# A REALIST EVALUATION OF FACULTY DEVELOPMENT IN UK MEDICAL SCHOOLS

By

**Olanrewaju O. Sorinola**

A thesis submitted to

The University of Warwick

For the degree of PhD

Doctor of Philosophy

Warwick Medical School

The University of Warwick

Aug 2014

# Table of Contents

# Tables

# Figures

# Boxes

# Acknowledgments

My special thanks and appreciation go to my supervisors Jill Thistlethwaite, David Davies and Ed Peile, who have supported and encouraged me throughout this PhD journey especially during times of self-doubt. I also need to thank my colleagues who have tolerated my need to discuss realist research and proof read my work with ingenuity and patience. I also wish to express my gratitude and thanks to all the educators and faculty development coordinators who contributed to this thesis. It would not have been possible without their participation. Finally, I would like to dedicate the thesis to my family for without their support, forbearance and encouragement this thesis and my PhD journey would never have been completed.

# Declaration

This is to confirm that this thesis is my own work and that the thesis has not been submitted for a degree at another university. In the appendix I have included my publications related to this thesis.

# Summary


## Background

The central mission of every medical school is education, however most teachers have not received formal teacher training. Hence there is an expectation that faculty development (FD) with planned programmes to prepare faculty members for their role in teaching will be integral to medical school life. The widespread investment in FD is predicated on the belief that it enhances the effectiveness of teaching but the evidence is limited.

## Aim

To carry out an in-depth exploration of FD in UK medical schools and evaluate the effectiveness of FD on teaching. The realist framework with its principle of explanatory causation was chosen to find out what works for whom in FD in what context, and why. This was an innovative use of the model in educational research.

## Methods

A detailed literature review was carried out and combined with the researchers experience and insight of FD to develop eight realist hypotheses in the form of contexts, mechanisms and outcomes (*CMO*). Data that would support, modify or challenge the hypotheses were then collected in three phases. Phase I was a review and scoring of data on medical school FD webpages using a webpage scoring index. Phase II was observation and informal interviews of two cohorts of educators attending a FD course followed by detailed interviews of 12 educators (six from each cohort) six months later. Phase III was interview of FD coordinators and educators at eight medical schools. Quantitative and qualitative data were collected.

**Data analysis**

Quantitative data were analysed by descriptive methods, slope diagrams, bi-axial constructs and statistical analysis with Fisher's exact test. Qualitative data were analysed by categorising and connecting strategies followed by summarising the relevant data under each hypothesis to check if the hypothesis was supported, modified or refuted.

**Findings**

The literature review revealed a paucity of publications on FD in the UK, however the 30 medical schools webpages reviewed showed data on FD activities. Data from all three phases were used to identify and explain the contexts (participatory approach/reflective practice, needs identification, supportive setting and standardization of medical teaching) that facilitated FD mechanisms (engagement, motivation, positive perception and professionalization).These led to the outcomes of improved confidence, competence, credibility and career progression.

**Conclusion**

Four realist theories of FD were identified (engagement, motivation, positive perception and professionalization) confirming the effectiveness of FD both in the short and long term and important recommendations for all FD stakeholders (educators, FD developers, universities and policy makers). In addition, I also made recommendations for future medical education researchers who choose the realist evaluation model.

# Glossary

| | |
|---|---|
| Beliefs | A belief is part of system that includes our values and attitudes, our personal knowledge, experiences, opinions, prejudices, morals, and other interpretive perceptions of the social world. |
| BEME | Best Evidence Medical Education |
| CMO | Contexts: Consists of the broader historical, cultural, economic, geographical, and structural factors that exist at the time of the initiative.<br>Mechanisms: They are the agents of change. They describe how the structures and resources embedded in a programme influence the reasoning and behaviour of the programme subjects. They are processes operating within an intervention that describes how the 'human components' use the resources available to them.<br>Outcomes: Follow from mechanisms acting in contexts and provide the key evidence for the realist evaluator to mount, monitor, or modify a programme. |
| Deanery | A regional body in the UK overseeing education in that region. Has now been replaced by Local Education and Training Boards (LETB). |
| ECE | Essentials of Clinical Education course at Warwick Medical School (WMS) |
| Ethnography | An approach to the description and interpretation of the culture and social structure of a group. It is an exploration and understanding of social settings and processes. Originally focused on primitive and exotic cultures, but now used more generally. |
| GMC | General Medical Council in UK regulates doctors |
| HEA | Higher Education Academy |
| HESA | Higher Education Statistics Agency |
| LETB | Local Education and Training Boards: see under deanery |
| Mixed Methods | Research that integrates quantitative and qualitative research within a single project such that the data derived are mutually illuminating |
| NSS | National Student Survey |
| PGCE | Post Graduate Certificate in Education |
| SHA | Strategic Health Authority |
| Themes | Theme (as in thematic analysis) represents a patterned response or meaning within the data set. It is the unit of analysis where the interpretative analysis of the data occurs in relation to the phenomenon being examined |
| UKPSF | UK Professional Standards Framework |
| Values | The importance we attribute to oneself, another person, thing or idea. |
| WFME | World Federation of Medical Education |

# CHAPTER 1
# INTRODUCTION

REALIST EVALUATION OF FACULTY DEVELOPMENT IN UK MEDICAL SCHOOLS

## 1 Background

The central mission of every medical school is education, however most medical teachers have not received formal teacher training (McLean et al, 2008; Steinert, 2010a; 2010b). Doctors are experts in *what* they teach but most have had little or no training in *how* to teach, hence the suggested need for faculty development (FD) with its principal purpose of improving teaching and, ultimately, patient care (Boelen, 1999; Branch et al, 1997; MacDougall & Drummond, 2005). Medical faculty development is said to represent an investment in human capital by nurturing and developing staff. The suggested benefits to individual faculty are improved vitality and growth that help to sustain them in their careers; institutions supposedly receive a return in the form of quality improvement over time while the profession receives a return through improved training for the next generation of doctors and improved patient care (McLean et al, 2008). Hence, there is an expectation that FD will be integral to medical school life.

To focus this chapter, I will start by giving a clear definition of FD, followed by a description of the influences on my development of the research and how the research question evolved. I will then give a brief description of realist evaluation and finish the chapter with an outline of the thesis.

## 1.1 Definitions

To provide a background to FD, I will define the word faculty. In North American English, the word *faculty* is used as a collective noun for the academic staff of a university. However, medical student education in the UK is carried out by various groups, for example: academic staff, clinical academics, clinical National Health Service (NHS) staff and nursing staff. Therefore, in this thesis, the term *faculty* is used in an inclusive way and refers to all individuals who are involved in the teaching and supervision of medical students, at all levels, in a wide range of contexts (e.g. in the classroom, at the bedside, in the outpatient clinic) and settings (e.g. the university, the hospital and the community). Similarly, I use an inclusive definition of *educator* to refer to individuals with specialist knowledge in the theory and practice of education who provides instruction and teaching. However, for the sake of simplicity, I will use both terms interchangeably in the thesis to refer to the same group of individuals.

Next, I define FD itself. While there have been many definitions of FD over the decades, I favour Bland and colleagues' (1990) definition of FD as a planned programme to prepare institutions and faculty members for their roles in the areas of teaching, research, administration and career management. FD in medicine has taken place since the late 1970s, stimulated by the growing demand for more innovative teaching (Herrmann et al, 2007). While there has been continued emphasis on teaching in FD there is now also a broader coverage of other faculty roles such as organisational and leadership development (McLean et al, 2008).

Lastly, I consider evaluation. Since my research is going to be evaluative, it is important to have a good working definition of evaluation. Hence, I consider evaluation along the lines suggested by Patton. According to Patton (2011), evaluation is ultimately about reality testing: getting real about what's going on, what's being achieved, examining what's working and what's not working. Patton (2002) defined evaluation as,

> ***"The systematic collection of information about the activities, characteristics, and outcomes of a programme to make judgments about the programme, improve programme effectiveness, and / or inform decisions about future programming." (Patton, 2002, p. 10)***

In the next section, I'll give a brief description of how the research evolved and the influences on the development of the research inquiry including my own background / role as the researcher and my beliefs and views on FD.

## 1.2 Development of the research inquiry

I am a clinician, a consultant in obstetrics and gynaecology who has worked for many years in educational environments and has research experience. It is through my working relationship with educational colleagues that this research study developed. Before it began, I was working with the West Midlands deanery (see glossary, pg. viii) on 'Training the Trainers' (TTT) courses at the postgraduate level. This subsequently led to my involvement in the Evidence Based Medicine (EBM) Unity project[1]. Funded by the European Union Leonardo da Vinci Vocational Training Action Programme, we developed a web based course on how to impart clinically relevant EBM teaching in various clinical environments with input from experienced

[1] EBM Unity project: http://ebm-unity.pc.unicatt.it/

EBM teachers, clinicians and educationalists from institutions in seven European countries (Thangaratinam et al, 2009; Weberschock et al, 2013). These two projects, (TTT and EBM) led me to thinking about FD in medical schools, the teaching of the future generation of doctors and how I might research this area. However, it was not until I was able to study for a doctorate degree in 2009 that these ideas about research became a reality. Nonetheless, the research project has to be understood as a constantly evolving project as I developed more critical ideas about the conduct of the research as time went on. As a background to understanding the framework for the research project, I now give a description of my role as the researcher and my perspective of FD.

### 1.2.1 Role as researcher

My involvement with postgraduate deanery TTT and the EBM European projects mentioned above exposed me to medical educators' lack of confidence and performance in teaching. In my mind, there is no doubt that medical educators need proper teaching skills before they can effectively help learners achieve the intended outcomes. Therefore, my purpose for carrying out this research project is based in my belief about the importance of teaching skills in achieving the intended learning outcomes. This belief originated from the pioneering work of Spady (1988) on outcome based education (OBE). He said,

> ***"Outcome based education means organising for results: basing what we do instructionally on the outcomes we want to achieve." (Spady, 1988, p. 9).***

Other authors (discussed in Chapter 2) have expanded on this by defining OBE as a way of designing, developing, delivering and documenting instruction in terms of its

intended goals and outcomes i.e. the product defines the process (Harden, 2002; Harden et al, 1999; Schwarz & Wojtczak, 2002). Leinster (2002) has drawn attention to the need to rethink how we educate doctors, taking into account among other things the changing roles of healthcare professionals, the need to be able to assimilate, evaluate and use new information and the importance of attitudes and communication skills. Furthermore, national and international reports (General Medical Council, 2002; 2010; Quality Assurance Agency for Higher Education, 2000; Institute for International Medical Education, 2002), have all emphasised the importance of OBE and said clarification of the learning outcomes in medical education helps teachers to decide what they should teach and assess, and students what they are expected to learn.

I am aware of my positive belief that training teachers to teach can be a way to produce good educational outcomes for students and, although this motivated me to initiate the research, I did attempt to control for confirmatory bias throughout the study. My stance is that I should identify my beliefs and values as mine, as denial of motives in an attempt to be objective and bias-free is not good as it obscures the influence of motives, assumptions and beliefs on the research conclusions. Moreover, it can cut off research from valuable insights, questions and guidance (Maxwell, 2012).

### 1.2.2 Faculty development – My perspective

My view is that at the heart of the university / medical school are its faculty members, the men and women who devote their lives to the teaching and service mission of the medical school and the health service. Many would agree that the quality of an

institution, though influenced by many factors, is related most closely to the work of the faculty. Their expertise, commitment, and energy directly shape the experiences of students and the impact of the institution on the broader community. Hence in my opinion, I think FD is important, as it can be an effective way to address the teaching needs of medical educators. This view seems to be shared nationally and internationally as the recent GMC (2012) document on the recognition and approval of trainers in the UK has now recommended a mandatory provision of FD activities in UK medical schools. Globally, the World Federation for Medical Education (WFME) has recommended that medical teachers professionalise their practice through teaching qualifications in conjunction with a medical school policy to include teacher training, development, appraisal and reward (WFME, 2000). Similarly, the World Health Organization (WHO) in its guideline on 'transforming and scaling up health professionals' education and training' (WHO, 2013) recommended that health professionals' education and training institutions should consider designing and implementing programmes for faculty and teaching staff relevant to the evolving health-care needs of their communities. It also emphasised the importance for governments, funders and accrediting bodies to consider supporting the implementation of higher education policies for mandatory FD programmes.

However, I have to admit that, although the call for investment in FD is predicated on the belief that it enhances the effectiveness of teaching and student learning, evidence of long-term impact is still limited (Steinert et al, 2006) and the fact remains that there are only a few generalisable evaluations of FD available to help FD developers (Mcleod & Steinert, 2010). Therefore, despite the clamour for institutions to use FD as a lever for change (Sorcinelli et al, 2006), the argument remains in

balance. According to Kirkpatrick (1994), four conditions are necessary for change to occur: the person must have the desire to change, knowledge of what to do and how to do it, a supportive work environment, and rewards for changing. Interestingly, the first two elements of change can potentially be achieved through FD activities, the last two cannot, and yet it is at the level of work environment that we expect change to occur. Steinert and colleagues' (2006) agreed with this and suggested that examining organisational characteristics, as well as the impact of FD on the organisation, is critical. Therefore, the question about the effectiveness of FD still remains unanswered and I explore this further in the next sections under the research question and the purpose of the study.

## 1.3 Research question

Following from the above, I constructed my research question as an outcome based question from my views on OBE and its influence on my thinking. My initial research question was,

> ***Is there evidence that faculty development is effective for training medical educators and does it lead to good educational outcomes for participants?***

However, this outcome based question could simply be answered yes or no and did not satisfy my desire to carry out a deep exploration of how FD interventions improve educational outcomes. I realised that a focus on answering the question 'is FD effective?' is not that useful as it bypasses the importance of context and environment on the success of educational initiatives. Therefore, I looked to realist evaluation as a more useful framework as it answers 'what works for whom in FD and why?', rather than 'does FD work?' Realism utilises contextual thinking to

address the issues of ‘for whom’ and ‘in what circumstances’ a programme will work, as it is axiomatic that certain contexts will be supportive to the intervention and some will not (Pawson & Tilley, 1997; 2004). A realist evaluation will prove more valuable in explaining if FD is fit for purpose, what works (or doesn’t), in what context, and why. Hence my research question became,

> ***Does evidence provided by the stakeholders suggest in what circumstances (C) faculty development programmes (M) can be effective for training medical educators and lead to good educational outcomes (O)?***

## 1.4 Aim / Purpose of study

The purpose of this study is to carry out an in-depth exploration of FD in UK medical schools and explore, using the realist framework, what works, for whom, in what context, and why. I aim to map and develop an understanding of the various FD interventions in use, and seek the views of educators on their experiences / perceptions of FD in relation to attributes such as usefulness, relevance, effectiveness and outcome. While the immediate focus of my research is to explore how medical educators engage with FD activities and the outcomes of this, I will also be able to give rich descriptions of FD having gained an understanding of the nature and effectiveness of the interventions from the educators’ perspectives, and relate my findings to the intended outcomes of the interventions. This is an exploration that has not been carried out before and is particularly important in the context of major investment in time and resources for FD. My aim is to recommend strategies for FD to enhance acceptability, efficacy and utility to participants once the contexts activating the causal mechanisms and leading to particular outcomes are understood. I will be able to produce pragmatic guidance that may be used by FD

developers to optimise the design of their interventions and by potential learners to evaluate whether a particular FD activity is right for them.

## 1.5 Evaluation research

I decided that an evaluation (as defined in section 1.1) research would be the most appropriate way to answer my research question. Various authors have described different types of evaluations including *process evaluation* to answer question like 'What do participants experience in the programme?', *outcome evaluation* to answer questions such as, 'To what extent are desired participant outcomes being attained?' and *developmental evaluation* focused on developing innovation using the complexity design (Patton, 1997; 2011; Houlden & Collier, 1999). Despite the terminologies, Rallis and Rossman (2003) and Bakken (2002) highlighted that evaluation of educational programmmes should consist of three major components: evidence description, criteria for comparison, and prediction / judgement. Description of evidence focuses on patterned observations or participants responses of attributes and details that assess quality. Criteria for comparison explore how the programme measures up to a standard or ideal while prediction provides recommendations for change, if needed, and how those changes might be implemented. Based on this information, I decided on evaluation research as appropriate for my study and the next task was to identify the type of evaluation.

### 1.5.1 Realist evaluation

For reasons that will be discussed in detail in Chapter 3, I chose the realist evaluation, recognising that I would need to apply this framework innovatively as it has not been commonly used in medical education research (Wong et al, 2012). A

brief outline of aspects of realist evaluation are included here because these govern much of the discussion in Chapter 2's literature review but the full explanation of the realist evaluation and the justification for its use are detailed in Chapter 3.

The philosophy of realism combines aspects of positivism and interpretivism and is particularly appropriate for research in the social world, which often involves the exploration of intricate phenomena (such as FD), which can be understood and explained in different ways (Maxwell, 2012). To make it easier for the reader to follow the discourse in Chapters 2 and 3, I will outline the realist researcher's interpretation of 'theories' since it is the task of the realist researcher to construct theories that might explain the social phenomenon under study and then to test those theories using rational criteria (Robson, 2011).

The term 'theory' has a certain meaning in realist research and, where it is necessary to distinguish it from the more general meaning of 'theory', I will use the term realist theory. Also, I will use the term 'Hypotheses of the Inquiry' in order to identify the realist ideas that I develop about FD during the study. The terms context, mechanism and outcome are commonly used in the English language but, within realism, they have quite specific meanings. In order to avoid confusion, for the most part I will use the terms only with their realist meaning throughout the thesis.

The theories and hypotheses used in realist research need to be constructed in terms of *Contexts, Mechanisms,* and *Outcomes* (CMO) and the researcher seeks to explain the connections between mechanisms and outcomes as well as the influence of contexts. Pawson and Tilley (1997) developed a model of realist evaluation for

use in the social sciences, which can be adapted for use in medical education. When I interpret this model for FD, the mechanisms are the structures of FD that emerge as important for learning (e.g. motivation to learn and gain from the FD, engagement generated by the FD), the outcomes are what are expected from the FD programme (e.g. improvement in pedagogical skills) and the contexts are not only the geographical location of the FD (e.g. the medical schools) but also wider issues such as government initiatives and medical council policies. The first task of the evaluation is, therefore, using the literature review and other data, to identify possible contexts, mechanisms and outcomes (CMO – glossary, pg. viii) and construct initial hypotheses that may explain the workings of FD programmes.

Finally, as a realist researcher, I will use a collaborative approach with the stakeholders, explaining the overall conceptual structure of the study to them so that they have a clear understanding of the purpose of the research task as well as listening to their views on my research hypotheses (see Chapter 4). My view is that the research project should be ‘with’ rather than ‘on’ the stakeholders since they will have much to contribute to process of the research. They can provide inspiration, criticism and consolidation for the study and by ensuring that the research process involved an open exchange of views, I hope to achieve valid and useful results (Reason, 1999). In the next section, I will outline the thesis structure and give a brief description of each chapter.

## 1.6 Outline of the thesis

The thesis is structured as follows:

*Chapter 2 Literature review: Creating the framework for a realist evaluation*

The literature review traces FD back to its historical root and various educational paradigms that have influenced FD. In addition, the literature is interpreted from a realist viewpoint to identify contexts that are likely to be barriers or facilitators of FD, possible mechanisms, and previously reported outcomes. These are later used to contribute to the development of the hypotheses described in Chapter 3.

*Chapter 3 Theoretical Framework: The rationale for using a realist evaluation*

In this chapter, the use of realism will be justified through a discussion of its underlying philosophical principles. The model of realist evaluation will be scrutinised and developed in an innovative way for use in education. I will also describe how the hypotheses of the inquiry, which underpin the process of the research, are developed, using information from a variety of sources.

*Chapter 4: Aligning Methods to Methodology*

Having identified the theoretical framework, in this chapter I will articulate the strategy for the stakeholders needed to supply the information necessary to answer the research question. I will also discuss the data collection and data analysis strategy for each phase of the research, the rationale behind the choices and the limitations of each approach.

*Chapter 5: Methods*

The methods chapter describes the three phases of the research study in detail.

Phase I: A review of FD webpages on UK medical schools websites

Phase II: Observation of FD sessions / interviews of FD coordinator and participants at Warwick Medical School (WMS) FD course

Phase III: Interviews of FD coordinators and medical educators at eight other UK medical schools.

The hypotheses of the inquiry provide possible explanations for the connections between contexts, mechanisms and outcomes and it is the purpose of these phases of the research project to collect a range of information from stakeholders that will support, modify or invalidate the hypotheses.

*Chapter 6: Phase I Findings*

The findings from the webpage review will be discussed in detail in and a realist interpretation of the possible contextual factors (explicit and implicit) influencing FD provision in the various UK medical schools will be explored. There will be discussion of the categorical data derived in this phase to be used for purposive sampling in phase III.

*Chapter 7: Phase II Findings*

My findings from experiencing a FD activity with two groups of stakeholders are explained. As realists also recognise that there is a need to interpret meaningful actions, my understanding of the views and intentions of the participants will be

discussed, as this is integral to understanding the FD programme. Both quantitative and qualitative data will be presented.

*Chapter 8: Phase III Findings*

I will discuss in detail the findings from the interviews of the FD coordinators and medical educators from the eight UK medical schools. I will also explain how causal mechanisms depend on context and the outcomes reported by the stakeholders.

*Chapter 9: Discussion of Findings*

In the penultimate chapter, I will review the hypotheses of the inquiry using all the information from across the research study. The discussion will focus on whether the hypotheses were validated, disproved or modified as well as a discussion of the validity and reliability of the findings. The emergent realist *CMO* theories will be identified and stakeholders' views on FD will be discussed. The process of realist evaluation will be scrutinised and the successes and difficulties in using the model and the implications for its use in educational research will be discussed.

*Chapter 10: Conclusions and Recommendations*

In the final chapter, I will give my concluding views about the effectiveness of FD and make recommendations. I will also discuss and make recommendations on how this innovative use of realist evaluation can be further developed for educational research. Finally, I will discuss my own developmental journey as realist researcher.

# CHAPTER 2

# LITERATURE REVIEW

## 2 Introduction

In Chapter 1, I explained how the first task of realist evaluation is to use the literature review to identify possible contexts, mechanisms and outcomes and construct initial hypotheses that may explain the workings of FD programmes. In this chapter, I discuss how I conceived the literature review as a funnel, which compresses the available information on FD into a framework for a realist evaluation. I start off by defining the boundaries of the study through clarification of the various terminologies of FD used in different countries. This not only drew a circumference around the subject matter of the research but also explained some of the terms that are used throughout the thesis. From this emerged, what I termed, the core of the literature review which is a consideration of the evolution, content and evaluation of FD examining the fundamental issue of whether FD leads to better outcomes for education professionals and paying particular attention to UK medical schools compared with other countries. The last section of this chapter considers how the literature review interconnected with realism and shaped it as a framework for this study of FD. The contexts, mechanisms and outcomes (*CMO*) derived from the literature review are later used in construction of the 'Hypotheses of the Inquiry' which is discussed in Chapter 3.

### 2.1 Terminology

FD has already been defined in section 1.1 as part of faculty members' professional development and as suggested in Chapter 1, there is an expectation that FD will be

integral to medical school life. However, the status of FD varies greatly within differing national contexts: from established programmes in USA and Canada, to a developmental framework in the UK; and from compulsory engagement in Scandinavia, to limited FD activities in France (Saroyan & Frenay, 2010). The situation is further complicated by the diversity of terms used to describe this development (staff development, educational development and academic development) as well as the fact that some languages have no equivalent.

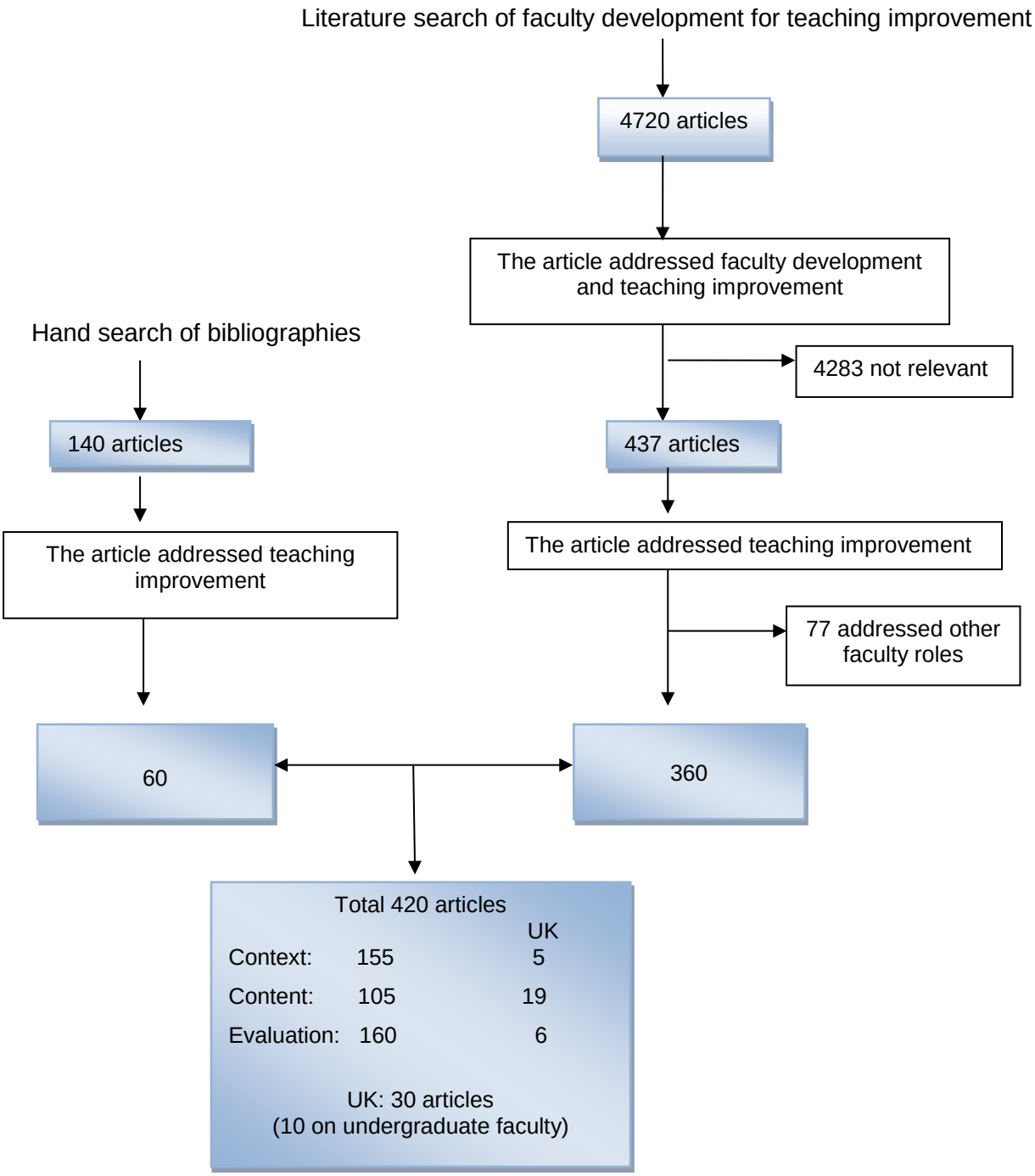
At the same time, the meaning of FD across various cultures is revealing (Steinert, 2012). For example, the Dutch term, *'docentprofessionalisering'*, loosely translates as the professionalization of teaching. This emphasis on professionalization, of both teachers and teaching, clearly aligns with a current focus on standards for teaching (GMC, 2006; 2012). The term is limited, however, in its emphasis on teaching (at the exclusion of other important faculty roles and tasks). In some ways, the French term, *'formation professorale'*, is more inclusive, as it is not restricted to teaching and refers to the 'formation' of the professorial role. The German term, *'Personal und Organisationsentwicklung'*, is also of interest, as it emphasises both individual and organisational development, another critical component of FD. In this literature review, the term FD is limited to teaching and covers interventions aimed at improving the educational skills of medical school educators while FD interventions solely designed to improve research, management skills and career development were excluded.

## 2.2 Search strategy

I carried out a literature search covering the period 1965 – Jan 2012 on Medline, ERIC, EMBASE, British Education Index, Australian Education Index and Teacher Reference Centre databases using the keywords: faculty development; medical faculty; staff development; and medical education. I manually searched major medical education journals (Appendix 1), proceedings of medical education conferences and experts' recommendations. The search was then repeated but limited to the UK. The databases were chosen as they cover the majority of the educational and medical databases and my search went back to 1965 to follow the historical evolution of FD, as it was in the second half of that decade that FD started to become prominent. All articles focusing on FD interventions for teaching were included but limited to publications in English. Articles focusing solely on other aspects of FD such as research, administrative skills and leadership skills were excluded. The search process is outlined in Figure 2.1.

Articles that focused on FD interventions in teaching and all study designs were included. Each article potentially meeting the inclusion criteria was screened for methodological quality based on previously published criteria from Best Evidence Medical Education (BEME) (Steinert et al. 2006). Full text papers of potentially relevant studies were assessed for relevance and inclusion (Figure 2.1). Data were extracted by adapting the BEME coding sheet which I had previously used in a systematic review of FD (Sorinola and Thistlethwaite, 2013).

**Figure 2.1: Literature review and selection of articles on FD for teaching**



## 2.3 Findings

From an initial 4720 abstracts, I found that 420 addressed medical faculty development on teaching. I read the 420 articles in full and to rationalise this large amount of information, I choose to group studies together under three main themes that represented different aspects of FD and also seemed able to provide possible contexts, mechanisms and outcomes for realist evaluation. Some studies did provide information on more than one theme. The themes were as follows:

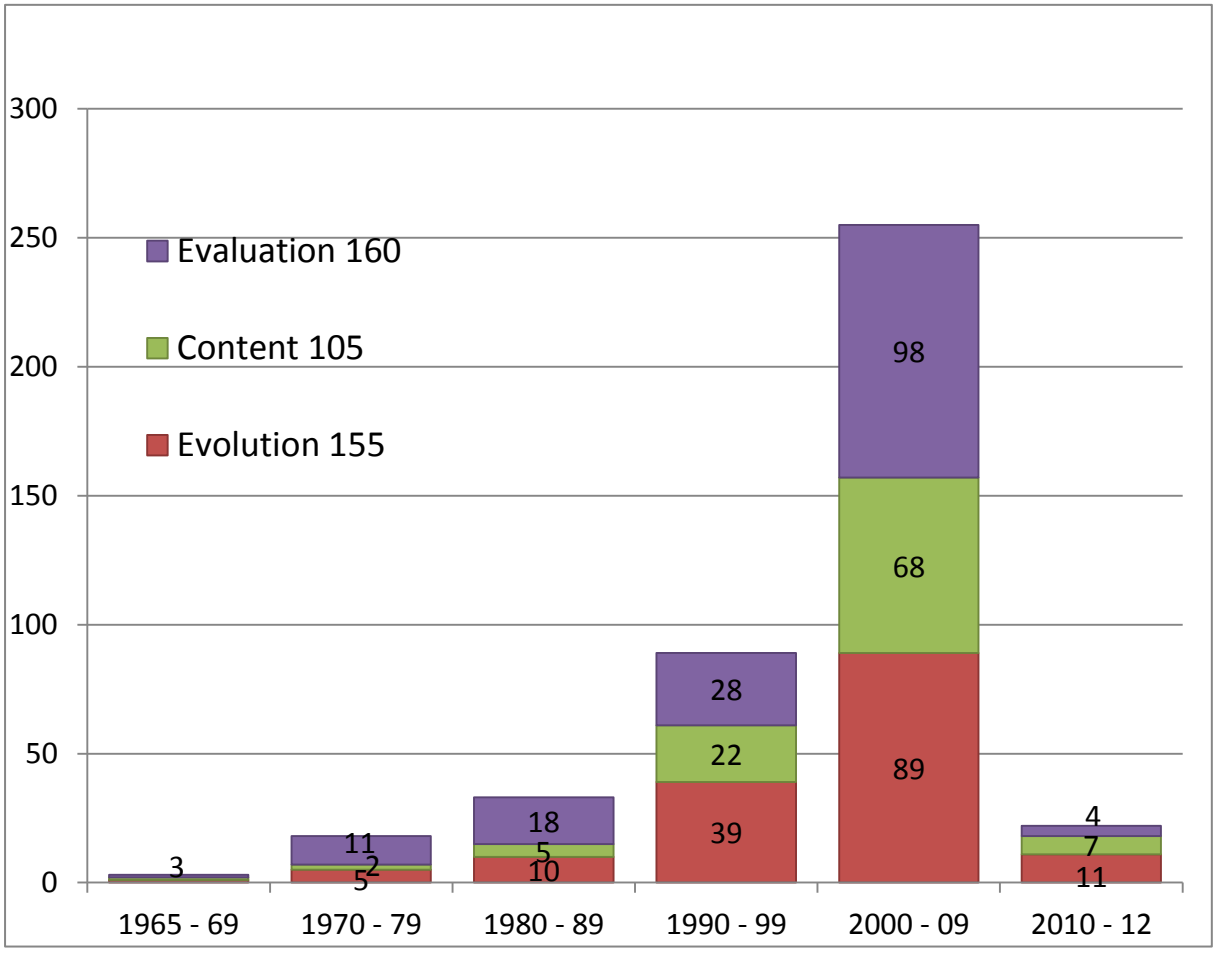
*Evolution*: articles that discussed the historical development, evolution, background, implementation, drivers and barriers affecting FD

*Content*: articles primarily focused on describing FD interventions or programmes on teaching even though they may have a small section on evaluation.

*Evaluation*: articles with detailed descriptions of FD programme evaluation.

The number of articles in each theme is shown in Fig 2.2, which also shows the spread of the articles over the decades. Over the last four decades (1970 – 2009), the numbers of articles have essentially doubled (or more than doubled) the preceding decade.

**Figure 2.2: FD articles by Themes and Decades**




It was suggested that the increase in the 90s was due to Boyer's (1990) identification of four categories of scholarship: discovery, integration, application, and teaching. According to Boyer, the scholarship of teaching, involves the capacity to effectively communicate one's own knowledge, skills, and beliefs; teaching becomes scholarship when it is made public, is available for peer review and critique so it can be reproduced and built on by other scholars. The increase in 2000s is thought to be due to the call for professionalization of teaching by various bodies (WFME, 2000; HEA, 2006; GMC, 2006; 2009). Professionalization in this thesis is defined as an attribute generating process that applies to individuals as well as institutions. It is a

process to attain and maintain professionality (compliance with agreed standards) as explained by Eitel et al. (2000) which I have discussed further in section 9.3.4. However, a careful look at the chart will suggest that by extrapolation, the number of published articles on FD has not grown at the same rate in the first three years (2010 – 2012) of the current decade. This might be because FD addressing teaching has been well covered and more articles are now been published on FD addressing other areas of educators development.

Grouping together different research studies was a problematic task since their designs were quite disparate and there were also methodological issues in some of the studies. The first issue was in the researchers' interpretation of FD, which, as noted above, has various synonyms and is sometimes confused with continuing professional development activities. The second was the multitude of instruments used to assess the impact of FD activities. They included self-evaluation by participants (Knight et al, 2007; Boerboom et al, 2009); self-evaluation in comparison with a control group (Branch et al, 2009; Dennick, 2003; Godfrey et al, 2004), questionnaires, focus groups, interviews, students' ratings, objective examinations, pre and post workshop evaluation (Pandachuck et al, 2004; Hewson & Copeland, 1999; Busari et al, 2006; Morrison et al, 2003), direct observation, portfolios and peer reviews (Hammersley-Fletcher & Orsmond, 2004; Gosling, 2002). As noted by Robson (2011), there is ambiguity in data collected using any of these instruments since respondents can interpret questions in different ways, individuals may have different agendas, and how qualitative data are analysed can depend upon the views of the researcher. Hence, although I have grouped the different research studies

under themes, this has to be understood within the limitations of combining different studies.

Bracketing off these limitations, I decided to combine the different studies in this way in order to identify possible contexts, mechanisms and outcomes for the realist evaluation. Pawson (2004), in his review of the quality of evidence in evidence-based policy, examined different kinds of research from a realist viewpoint. He made the distinction between the practice of research (the technical competence of the inquiry) and the progress of the inquiry and stated that research only progresses in so far as each investigation contributes a better set of explanatory propositions. For Pawson, the quality of a research study,

> ***"is not its technical competence as such, but whether its technical infrastructure provides good explanation." (Pawson, 2004, p. 3)***

I adopted a similar view in reviewing the literature on FD since I was looking for explanations, which might contribute to the structure of my research project. Fundamental to the design of a realist evaluation are the *CMO* hypotheses, which make explanatory propositions for the causal relationship between contexts, mechanisms and outcomes. Therefore, the aim of the literature review as part of developing the framework for a realist evaluation was to identify evidence and possible middle order theories (Chapter 3) that inform my hypotheses. However, for the sake of clarity, I have described below the initial findings under each theme before combining the data to develop the hypotheses.

## 2.4 Articles on evolution of FD

There were 155 articles covering the historical evolution, definitions of, drivers for and barriers to FD.

### 2.4.1 Historical evolution

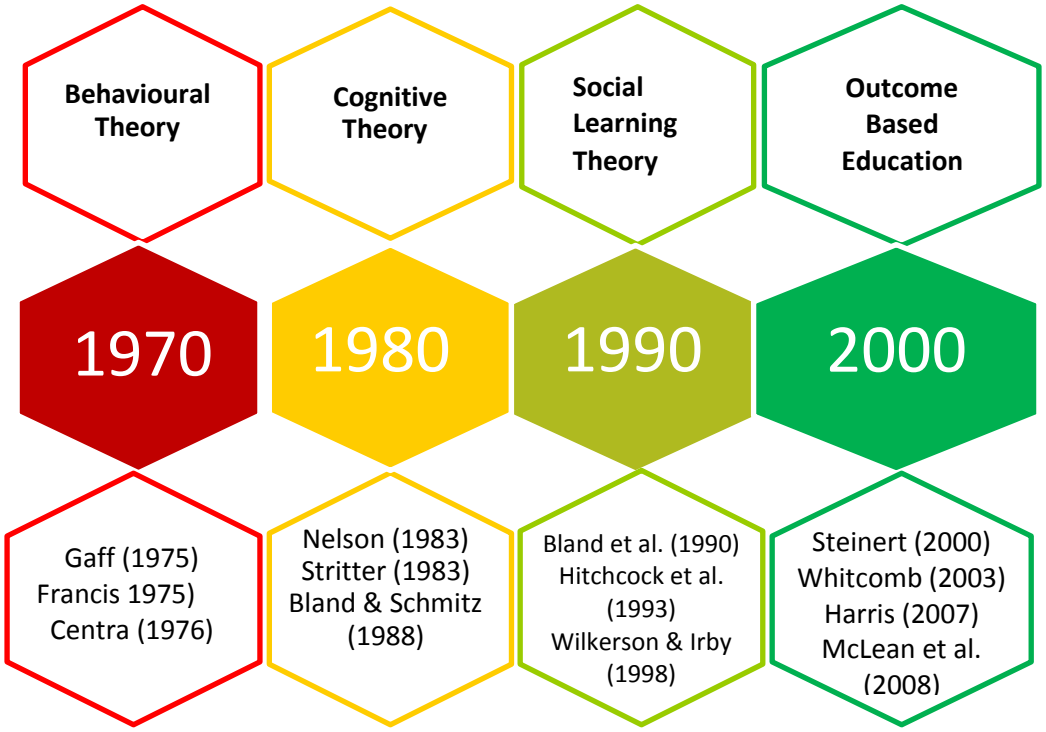
A brief history of FD is fundamental to understanding the divisions between education, clinical and research roles of teachers, since the development of FD highlights some of the tensions felt by teachers who are part of the health service yet work in medical education. Historically, FD in higher education meant developing expertise in one's discipline. The assumption for many years was that if one knew the subject one could teach it: content expertise translates into competent teachers (Hesketh et al, 2001; Seabrook, 2001). Therefore, most higher education institutions in the early 60s offered support for sabbatical leave, travel to meetings, conducting research and completing an advanced degree in the practitioner's specialty. The sabbatical leave, begun at Harvard in 1810, is the oldest form of FD in the USA and was the model until the early 1960s (Cochran, 1989). Usually, sabbaticals were competitive and given for projects that could not be completed on the home campus, requiring both travel and free time, hence their limitations (Eble & McKeachie, 1985). The emphasis was so much on expertise that Donald (1977) noted,

> ***"Because expertise and leadership in one's discipline are of paramount importance at this level of education, it has proved more difficult to examine the other important factors involved in the teaching process." (Donald, 1977, p. 11)***

Only in the second half of the 20th century has teaching in medical schools been acknowledged as a skill independent of content expertise. The theories underpinning

student learning played a major role in this shift in emphasis. They contributed to the evolution of FD with the various definitions of FD reflecting the changing theories in education (Figure 2.3).

**Figure 2.3: Evolution of FD and Learning Theories**




**(FD as defined by various authors – Appendix 2)**

In line with the behaviourist theory of the 70s, FD aimed to develop the attributes and competencies of the 'good' teacher; a good teacher was considered as someone who used various teaching aids, reinforced important concepts and communicated effectively (Gaff, 1975; Centra, 1973). Hence, the FD practices of that era included the description of best teaching practices and focused on modifying the behaviour of faculty members to achieve the desired effects (Gaff, 1975; Centra, 1976). The limitation of the theory was that it focused on that which was observable, disregarding anything occurring inside the mind of the learner.

Hence in the 1980s, the cognitive theories of learning, in which learning involves the active construction of meaning (Shulman, 1987; Bruer, 1993), began to compete with behavioural theories. Conceptual constructs and mental processes became the focus of interest requiring a metamorphosis of the teacher, from a didactic conveyer of knowledge to a facilitator of learning (Entwistle & Ramsden, 1983; Knowles, 1980). To make this transition teachers needed new skills, which required training, hence other definitions of FD (Appendix 2) were proposed (Stritter, 1983; Nelson, 1983; Bland & Schmitz, 1988).

Social learning theories emerged in the 90s. In the social constructionist view, learning was viewed as socialisation into a new knowledge community occurring through the learner's active participation in the community and the internalisation of socially constructed meaning (Good & Brophy, 1990; Jonassen, 1991). FD was therefore broadened to take cognisance of socialisation (Sheets & Schwenk, 1990; Bland et al, 1990), a view summarised by Wilkerson & Irby (1998) that becoming a medical teacher is more than the acquisition of appropriate skills; instead it is a socialisation process into the academic community comprising teachers, students and staff support personnel.

The year 2000 brought outcome-based education (OBE) into more focus. OBE owes much to the pioneering work of Spady (1988) as I discussed in section 1.2.1. Harden et al. (1999) expanded on this, describing OBE as,

> ***"An approach to education in which decisions about the curriculum are driven by the outcomes the learners should display by the end of the course i.e. product defines process" (Harden et al, 1999, p. 8).***

OBE can be summed up as a way of designing, developing, delivering and documenting instruction in terms of its intended goals and outcomes. It is a 'results orientated thinking' the opposite of 'input-based education' (where emphasis is on the educational process) and time based courses (Harden, 2002; Schwarz & Wojtczak, 2002). The various roles of the medical teacher, from clinical expert to mentor, supervisor, assessor and role model, were further expanded with important implications for FD. According to Steinert (2000), to keep pace with these changes, FD needed to broaden its focus by using diverse learning methods, underpinned by learning theories, and to rigorously evaluate the interventions. The wider roles expected of the teacher led Whitcomb (2003), Bligh (2005), Harris (2007) and McLean et al (2008) to redefine FD (Appendix 2).

### 2.4.2 Drivers for FD

Three main drivers for FD were identified by Gruppen et al. (2006): the need to sustain academic vitality, the changing nature of health care delivery, and public accountability. Vitality has been defined as,

> ***"The essential, yet, intangible, positive qualities of individuals and institutions that enable purposeful production."***
> ***(Clark & Lewis, 1985, p. 3; Bland & Schmitz, 1988, p. 43)***

Sustaining academic vitality is important as stress and burnout amongst medical teachers is common due to the increasing number of students, increasing workload, assessments, demands / expectations from different sources and implementation of new educational approaches without additional resources (Harden, 1999). The design and delivery of innovative FD activities has been suggested as one way of promoting and maintaining academic vitality and excellence (Gruppen et al, 2006;

Bland & Schmitz, 1988; Bligh, 2005). Similarly, FD has been facilitated by changes in medical education, which has evolved in order to respond to advances in medicine, alterations in healthcare delivery system, patient expectations, new learning technologies and developments in educational thinking (Harden & Laidlaw, 2012). The public also have changing expectations of medical practitioners with demands for medical training to be more authentic, responsive to societal needs, pressures for quality assurance, accountability, and standards in medical education (Leinster, 2002; McLean et al, 2008).

Another driver is the need for faculty retention as recruitment into academic institutions can be expensive (Waldman et al, 2004) hence FD is an important aspect of faculty recruitment and retention (Harden & Laidlaw, 2012). Another factor is medical teachers' lack of (or little) training for their teaching given that academic appointments in medical faculties are typically based on a combination of a professional qualification (clinical, biomedical or sociological) and research excellence (McLean et al, 2008; Higgs & McAllister, 2007a; 2007b). This lack of preparedness has become more acute with integrated curricula, new technologies (e.g. simulation and e-learning), older students and students from ever-increasing ethnic, cultural and socio-economic diversity (Gaff & Simpson, 1994; McLean et al, 2008). Other authors have emphasised the call for the professionalization of medical teaching. In the UK, the Higher Education Academy (HEA) has called for certification of teachers (Fry, 2006; HEA, 2006). The General Medical Council (GMC, 2006), UK stated,

*"If you have responsibilities for teaching, you must develop the skills, attitudes and practices of a competent teacher." (Good Medical Practice, GMC, 2006, p. 14)*

### 2.4.3 Barriers to FD

FD may be impeded by many factors including unsupportive leadership, financial / budgetary constraints, educators' attitudes / misconceptions (e.g. lack of motivation or unwillingness to acknowledge deficiencies in teaching ability), time pressure, and the lack of research on effective FD activities (Steinert, 2013; Hitchcock et al, 1993; Skeff et al, 1997a). Lack of institutional support is a key barrier as the value ascribed to FD is affected by the institutional culture, commitment, resource allocation and hence participation in FD (Gruppen et al, 2006; Simpson et al, 2006; Hitchcock et al, 1993; Ramani, 2006; Healey, 2000; Knight & Trowler, 2000; Richardson, 2005; Norton et al, 2005). While mission statements of most medical schools advertise teaching as a priority, it is research that is usually prioritised (Clark et al, 2004; Hitchcock et al, 1993). This is supported by Steinert et al's. (2009) finding of lack of appreciation, recognition or financial reward for teaching excellence.

As mentioned above, educators' attitudes and misconceptions can be a barrier. Educators may over or under estimate their teaching ability, may not perceive the benefits of training or fail to recognise a link between teacher training and teaching excellence (Skeff et al, 1997a). Another issue for medical teachers is time for FD activities because of the competing demands of clinical work and research (Steinert et al, 2009; Liben et al, 2005). According to Skeff et al. (1997a) teaching improvement requires time and faculty need to devote time to the pursuit of excellence in teaching.

## 2.5 Articles on content

There were 105 papers describing programmes, topics, formats and delivery of FD. The majority of the studies were carried out in North America and directed at improving teaching / scholarship of education. Examples include the Medical Education Scholars Programs (MESP) at the University of Michigan medical school, (Gruppen et al, 2003), the Teaching Scholars Programs (TSP) at the University of Washington medical school, (Robins et al, 2006), and faculty of medicine, McGill University, Canada (Steinert & McLeod, 2006). Other medical schools formed programmes called 'Academies' designed to share teaching techniques and experiences across schools and disciplines e.g. the 'Academy of Masters' collaborative between Harvard Medical School and the University of California, San Francisco (UCSF) (Thibault et al, 2003; Rider et al, 2002). These programmes share common goals (Fidler et al, 2007; Gruppen et al, 2006) including:

Promoting the scholarship of teaching
Enhancing teaching methods
Developing new teaching modalities
Promoting education research
Enhancing curriculum / assessments development
Promoting advising and mentoring
Promoting educational leadership skills

Again I have grouped these studies under the following sub-headings to summarise their key points.

### 2.5.1 Theory

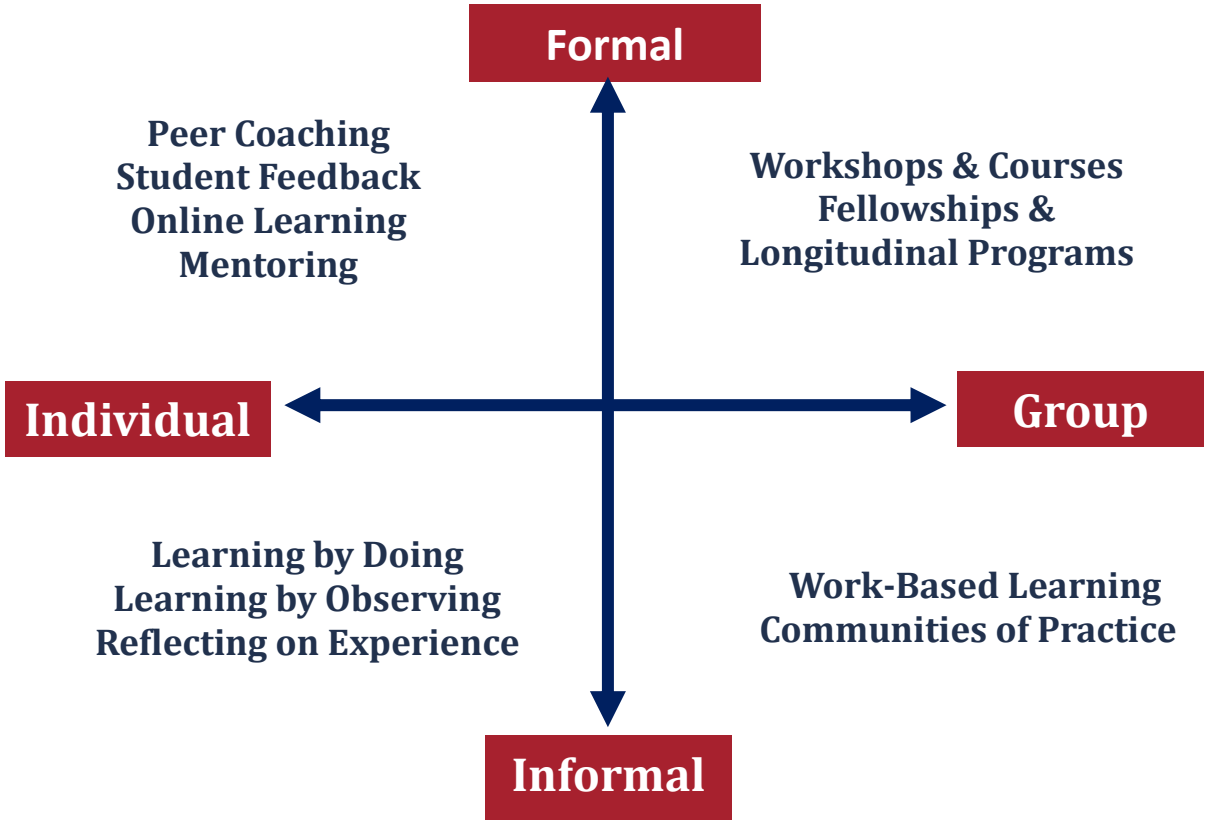
Some programmes were clear about their theoretical approach to learning but the majority were not grounded in a theoretical framework. Commonly described

strategies were adult education learning (McLaughlin, 2005), active learning (Frohna et al, 2006), self-directed learning, experiential / reflective learning (Amin et al, 2006), and conceptual change theory in which educators were helped to develop or change their existing educational ideas into more elaborate concepts of teaching e.g. from a teacher-centred to a learner-centred conception of teaching (Hewson, 2000; Light & Calkins, 2008; Pratt, 1992). Another example is the 'paired peer' learning model at the Boston University School of Medicine, in which participants partner to share their experiences and receive iterative cycles of teaching, debriefing, and planning (Orlander et al, 2000). The general consensus from the literature was that FD needed to be grounded in a theoretical framework, which should be made explicit (Steinert et al, 2012).

### 2.5.2 Types of FD activities / Instructional methods

FD tends to be delivered as short courses, workshops, seminars or accredited university awards: postgraduate certificates, diplomas and masters (Amin et al, 2006; Hewson, 2000). More innovative approaches include longitudinal programmes: faculty commit a proportion of their time on a regular basis over one to two years to develop knowledge and skills leading to advanced degrees or fellowships (Steinert et al, 2003; Steinert & McLeod, 2006; Bland et al, 1988); and hybrid programmes – a combination of workshops / seminars and online learning (Fidler et al, 2007). Figure 2.4 summarises types FD activities along two dimensions, from formal to informal and from individual to group activities. These FD activities happen to varying extent; workshops and courses were the most commonly utilised FD intervention while mentorship and communities of practice were the least utilised (Steinert, 2010a).

**Figure 2.4: Types of Faculty Development Activities**




(Adapted with permission from Yvonne Steinert's presentation at the 1st International Conference on Faculty Development in the Health Professions, Canada, 2011)

FD programmes ranged in duration: some as short as one hour and others as long as three years. Workshops ranged from three hours to one week, seminar series, occurring over time, were 12 hours to 28 days and short courses ranged from seven to 8 days. Fellowships ranged from one to three years; one of the longest was a three year long TSP at the, Carver College of Medicine (Iowa) leading to certification (Rosenbaum et al, 2006). Regarding instructional methods, most interventions used a multimodal approach such as small group discussions, workshops, interactive exercises, role play, simulations, videotaped teaching review, case based workshops, web-based learning and mentoring (Lang et al, 2000; Newland et al, 2003; Steinert et al, 2006; Boucher et al, 2006).

### 2.5.3 Topics / Focus of intervention

As a result of the selection criteria, all papers focused primarily on teaching improvement, but some also addressed organisational, leadership and research skills, personal / career development as shown in the four categories in Table 2.1.

**Table 2.1: Topics Covered in Faculty Development**

| Instructional | Professional |
|---|---|
| Philosophy / Theories of education<br>Concepts in adult education<br>Scholarship of teaching<br>Clinical teaching models<br>Writing objectives and designing instruction<br>Transforming the learning environment<br>Assessment and feedback<br>Formal presentation skills<br>Choosing educational strategies | Curriculum design / Needs assessment<br>Teaching portfolios<br>Mentoring<br>Peer observation and evaluation<br>Programme evaluation<br>Self-evaluation<br>Leadership / Career progression |
| **Research** | **Organisational** |
| Educational research<br>Research design<br>Ethics in research<br>Research methods / Data analysis<br>Literature searches on education topics | Time management<br>Planning and organising<br>Teamwork<br>Project management<br>Information Technology |

### 2.5.4 Participation

Most programmes were voluntary and available to all, but a few programmes, especially the longitudinal programmes, were competitive entry. Competitive entry requirements included an application form, a curriculum vitae, a letter of recommendation from the dean, student evaluation reports and interviews. Examples were the TSPs at Iowa Carver College of Medicine, USA, and the faculty of medicine, McGill university, Canada (Rosenbaum et al, 2006; Steinert & McLeod, 2006). Interestingly, no FD study reported compulsory participation or contractual requirement at the time of the literature review. However, as mentioned earlier the

recent GMC (2012) document on the recognition and approval of trainers in the UK has now brought a mandatory component to FD provision in UK medical schools.

While some programmes had only medical school faculty as participants, others were multidisciplinary involving other health sciences departments such as nursing, pharmacy, occupational therapy and communication sciences (Steinert & McLeod, 2006; Fidler et al, 2007). In addition, some interventions were university wide across different disciplines (Steinert et al, 2010; Postareff et al, 2007). However, only a few studies reported on participants progress as educational leaders and facilitators of FD (Steinert et al, 2001).

## 2.6 Articles on evaluation

I have already given a working definition of evaluation in section 1.1. However, for the review of evaluation studies reported in the literature, I found Mohanna et al's. (2004) description more apposite as they viewed evaluation as a systematic approach to the collection, analysis and interpretation of information about any aspect of the conceptualisation, design, implementation and utility of educational programmes. There were 160 articles focused on FD evaluation with the majority using the modified Kirkpatrick's (1994) model for evaluating outcomes, even though a few studies used other evaluation techniques such as Miller's (1990) developmental evaluation and the education value compass (Ogrinc et al, 1999).

### 2.6.1 Evaluation methods

Miller's (1990) developmental evaluation (assessment pyramid) for learners to demonstrate: what they know, that they know how, that they can show how, and that

they can do (equivalent to Kirkpatrick's levels 2 and 3) is useful when assessing the conjunction of knowledge and skills for a task. The model is extremely helpful when identifying the level and type of knowledge and skill that learners are achieving. The education value compass uses learner's knowledge, patient clinical outcomes, function, satisfaction, and costs as a framework for evaluating the educational endeavours that occur with clinical care. The education value compass emphasises a balanced set of measures including practice level changes and clinical outcomes for patients (Ogrinc et al, 1999). Each of these evaluation models is useful depending on the focus of the evaluation. However, these models fall short in a key component: none identify the depth of contextual information about the educational training that can be helpful when attempting to replicate findings to another setting. This lack of understanding of the contextual conditions of an educational intervention is a limitation (Ogrinc & Batalden, 2009).

As mentioned above, the Kirkpatrick's (1994) model was the most commonly used model for evaluation in the papers reviewed. The original Kirkpatrick's model described four levels of outcome: learners' reaction (to the educational experience); learning (which refers to changes in knowledge and skills); behaviour (which refers to changes in practice and the application of learning to practice); and results (which refers to change at the level of the learner and the organisation). Freeth et al. (2002) modified the model for their evaluation of interprofessional education (IPE). They modified level two 'learning' (distinguishing between change in attitude 2A and knowledge and skills 2B) and level four 'results' (distinguishing between outcomes that had an impact on service delivery 4A and those related to people 4B) to produce six categories. The modified Kirkpatrick's model was adopted by the Best Evidence

Medical Education (BEME) Collaboration[2] and has been used in other BEME reviews (Issenberg et al, 2005; Steinert et al, 2006). The model shown in Table 2.2 was further adapted by Steinert et al. (2006) to include students and residents (instead of patients) at level 4B in their BEME review.

**Table 2.2: Kirkpatrick's model for evaluating educational outcomes***

| LEVEL | CATEGORY | DESCRIPTOR |
|---|---|---|
| **Level 1** | **REACTION** | Participants' views on the learning experience, its organisation, presentation, content, teaching methods, and quality of instruction |
| **Level 2A** | **LEARNING** Change in attitudes | Changes in the attitudes or perceptions among participant groups towards teaching and learning |
| **Level 2B** | **LEARNING** Modification of knowledge or skills | For knowledge, this relates to the acquisition of concepts, procedures and principles; for skills, this relates to the acquisition of thinking / problem-solving, psychomotor and social skills |
| **Level 3** | **BEHAVIOUR** Change in behaviour | Documents the transfer of learning to the workplace or willingness of learners to apply new knowledge and skills |
| **Level 4A** | **RESULTS** Change in the system / organisational practice | Refers to wider changes in the organisation, attributable to the educational programme |
| **Level 4B** | **RESULTS** Change among the participants' students or residents. | Refers to improvement in student learning / performance as a direct result of the educational intervention |

*Kirkpatrick's model (1994) was modified by Freeth et al. (2002). The model shown here was adapted further by Steinert et al. (2006) to include students at level 4B in their BEME review.

The BEME evaluation approach has been widely adopted, in my view, largely because the notion of a hierarchy of evidence resonates strongly with two dominant themes in clinical medicine, namely: the 'evidence hierarchy' (from case studies at the bottom to statistical meta-analysis at the top) of evidence based medicine (EBM) and the valuing of 'hard' clinical outcomes over 'soft' outcomes in health services

[2] BEME Collaboration: An international group of individuals, universities, and organizations committed to synthesizing and disseminating the latest educational research findings in order to provide a basis for informed decision making. www.bemecollaboration.org

research, against which medical education research has been unfavourably compared (Todres et al, 2007; Stephenson et al, 2009).

While the Kirkpatrick model has many advocates, in my view the model has some limitations in evaluating educational interventions in general and FD in particular. There are a number of reasons for this. In his original work, Kirkpatrick (1967) asserted that these outcomes were not hierarchical and that the model is intended to provide a comprehensive evaluation that can inform policy and programme development. In a more recent book, Kirkpatrick (2009) explained how he arrived at the set of descriptors that are now widely used to evaluate educational interventions. He had observed that technical training could be evaluated by measuring learners' reactions, learning and behaviour, and their impact on the organisations for which the learners worked. Kirkpatrick's original purpose was to provide business managers with promptly identifiable and easy-to-measure outcomes in learners and the organisations for which they worked. Moreover, of his numerous references to successful applications of the model, none came from a field as complex as medical education, which differs from business in that it is required to equitably meet the needs of a whole array of beneficiaries such as, patients, students, practitioners and health care organisations. The model does not allow for the rich variety of outcomes that can be evaluated using qualitative as well as quantitative methodologies, nor explain how or why such outcomes are consequential to particular elements of complex interventions (Yardley & Dornan, 2012).

Yet, unless we understand how and why effects are consequential to particular elements or interactions, it will be difficult to refine education to maximise benefit.

Yardley and Dornan (2012) argued that where medical education really deviates from evidence based medicine (EBM) is in its recognition of a wide gap between the results of simple experiments and their applicability in 'real practice'. Context as well as process impacts on educational outcomes and they asked the question: Are outcomes necessarily more important than processes (which are not included in Kirkpatrick's levels)? Holton (1996) also criticised the model on the grounds that it lacked important attributes of a theory and lacked supportive evidence to indicate that lower-level outcomes are prerequisite to higher-level ones.

For FD evaluation specifically, there are a number of shortcomings in using the Kirkpatrick's model. First, there is little comparative evidence on which components of FD are most useful or whether one method (e.g. workshop) is more effective than another (e.g. longitudinal courses) as the model does not delve into such depths. According to Steinert et al (2006) example, workshops, though one of the most common FD activity, have been evaluated mostly at level one (reaction) and the impact on actual performance and the duration for which this effect was sustained was not usually assessed. Longitudinal FD activities, on the other hand, have the advantage of using methods such as reflective practice, skills workshop and repeated feedback, which may have a more lasting impact, but the long-term effects of these resource-intensive longitudinal FD have not been looked at in detail either (Knight et al, 2007). I accept that there may be various explanations for this, such as the large numbers of intervening variables arising during the follow-up period, which cannot be effectively controlled, such as personal attributes (e.g. change in teacher's status, seniority, educational experience) or process variables (e.g. curriculum change, modifications in class-size, differences between student cohorts). However,

it still leaves the question unanswered as to the superiority of one approach over the other. Some authors on the other hand, have suggested that it is the experience of FD rather than the particulars of length, content, or delivery that have lasting importance (Barlett & Rappaport, 2009). Whatever the argument, I am of the opinion that a realist approach which takes all these contextual issues into consideration will be able to provide explanations as to their influence or otherwise on the outcomes as it recognises that educational interventions take place in open, complex systems.

Second is the issue of participant characteristics. The issue of self-selection has been the subject of continuing debate as some researchers question whether educators who chose to participate in FD are already more motivated, committed or differed in professional or personal characteristics from non-participants (McGaghie et al, 1990; Notzer & Abramovitz, 2008). Other researchers, however, support the principle that it is more effective to teach the academic skills needed for success to educators who already have the idealism, motivation and commitment (Freeman et al, 1998). Again a realist approach to evaluation as opposed to Kirkpatrick levels should provide explanations of the influence of participants' characteristics on the educational outcome.

Lastly, as summarised by Teherani et al. (2001), the design, objectives, and participants vary between FD interventions, rendering research on FD context-dependent and making it a complex process to consistently measure one outcome from a variety of different interventions. Hence, another reason to choose a realist evaluation framework to explore the influence of contexts on mechanisms and

outcomes. In the section below, I have summarised the available publications on evaluation.

### 2.6.2 Non-systematic reviews

The key question in FD has always been: 'is there evidence of improved teaching effectiveness following a FD intervention?' Attempts to answer the question have been going on for decades. In the 80s, several authors reported that despite the growth in FD, evaluation of these initiatives were still a rare occurrence, usually consisting of short questionnaires on participants' satisfaction (Bland & Schmitz, 1986; Stritter, 1983; Sheets, 1985; Sheets & Henry, 1984; 1988). At the start of the 90s, Sheets & Schwenk (1990), reviewed the literature on FD for educators and made a similar observation, calling for more rigorous evaluations. In 1993, Hitchcock et al. reviewed FD literature and concluded that the concept of FD was evolving and expanding but the efficacy and outcomes needed better documentation. Reid et al. (1997), reviewed 24 papers (published 1980-1996) and concluded that despite some positive outcomes for fellowships, workshops and seminars, methodological weaknesses precluded definitive conclusions regarding the impact of FD. All the above reviews were limited by the lack of a systematic search of the literature and the absence of a pre-determined protocol for reviewing the articles, but nevertheless they all suggested that a more rigorous evaluation of FD was needed.

### 2.6.3 Systematic review

There has been only one systematic review of FD initiatives on teaching improvement. Steinert et al. (2006) used three databases to search between 1980 – 2002 and reviewed 53 articles (six randomised controlled trials and 47 quasi-

experimental studies), of which 31 used a pre-test / post-test design. They concluded that FD initiatives appeared highly valued by participants, leading to changes in learning and behaviour but changes in organisational practice and student learning were not frequently reported. The review summarised the characteristics of effective FD as using multiple instructional methods, experiential learning approach, peer / collegial support and adherence to principles of teaching and learning. However, limitations of the studies in the review included; absence of comparison groups, issues with robustness of the outcome measures such as over-reliance on self-reporting, use of Kirkpatrick level 1 reaction mostly, few measures of impact on learners and underuse of qualitative methods. In addition, other authors (Postareff et al, 2007) have criticised the review in that the majority of interventions included targeted didactic skills in the classroom and at the bedside with little attention to other teacher roles, such as organiser or developer of education.

Taken together, the conclusion from these publications was that there was a paucity of research on the long-term benefits of FD despite several decades of research even though there are some reported successes of FD interventions (Skeff et al, 1997b; Skeff et al, 1997a; Steinert, 2013; Steinert et al, 2006; Guskey, 2003; Prebble et al, 2004). Moreover, while efforts have been made to use various methods to evaluate the impact of FD, success has been limited by the overall lack of a good instrument for measuring teaching effectiveness. Consequently, no conclusions can be made as to the most effective FD intervention or the optimal time for reinforcement. Hence, while the ultimate aim of FD is to improve teaching practices, the quality of medical graduates and ultimately patient care, to date, the evidence for this impact remains weak (Steinert et al, 2006).

## 2.7 FD in the UK

In the UK, over the last two decades, there has been a growing consensus that educators need to be trained in educational methods. In 1991, the Committee of Vice Chancellors and Principals of the Universities of the UK called for more training in educational methods for all university teachers (Elton, 1991). A national inquiry into medical education identified the need for professional expertise in curriculum development, teaching methods, assessment and training of academic staff (Towle, 1991). The National Committee of Inquiry into Higher Education recommended accredited training for all academics at higher education institutions (Dearing, 1997). Furthermore, the GMC in its agenda for undergraduate medical education recommended that all medical graduates should have an awareness of the principles of teaching and learning (GMC, 2003; 2006; 2012). As previously mentioned, the HEA has also called for certification of teachers (including medical teachers) and emphasised the importance of teaching scholarship (Fry, 2006; HEA, 2006). So how is this currently interpreted in UK medical schools?

The UK has 33 medical schools (Table 2.3) employing 3358 full time equivalent clinical academics and many more non-clinical academics (Medical Schools Council, 2010a; 2010b) and yet, surprisingly, my literature search revealed only 30 published papers between 1965 and 2012 addressing FD activities in the UK (Figure 2.1).

## Table 2.3: UK Medical Schools by Country / Region

| | |
|---|---|
| **England** | |
| | **East Anglia** |
| 1 | Cambridge (University of), School of Clinical Medicine |
| 2 | University of East Anglia |
| | **Greater London** |
| 3 | Barts and The London School of Medical and Dentistry |
| 4 | King's College London School of Medicine (Guy's King's College St Thomas' Hospital) |
| 5 | Imperial College School of Medicine, London |
| 6 | London School of Hygiene and Tropical Medicine (Postgraduate Medical School) |
| 7 | St George's, University of London |
| 8 | University College London, University College Medical School |
| | **Midlands** |
| 9 | Birmingham (University of), School of Medicine |
| 10 | Keele University, School of Medicine |
| 11 | Leicester (University of), Leicester Medical School |
| 12 | Warwick (University of), Warwick Medical School |
| 13 | Nottingham (The University of), Faculty of Medicine and Health Sciences |
| | **North East** |
| 14 | Durham (University of), Queens Campus, Stockton, Phase 1 Medicine |
| 15 | Hull York Medical School |
| 16 | Leeds (University of), School of Medicine |
| 17 | Newcastle (University of), Newcastle Biomedicine, The Medical School |
| 18 | Sheffield (The University of), School of Medicine |
| | **North West** |
| 19 | Liverpool (University of), Faculty of Health and Life Sciences |
| 20 | Manchester (University of), Faculty of Medicine and Human Sciences |
| | **South** |
| 21 | Brighton and Sussex Medical School |
| 22 | Oxford (University of) |
| 23 | Southampton (University of), School of Medicine |
| | **South West** |
| 24 | Bristol (University of), Faculty of Medicine |
| 25 | Peninsula Medical School |
| **Scotland** | |
| 26 | Aberdeen (University of), School of Medicine |
| 27 | Dundee (University of), Faculty of Medicine, Dentistry and Nursing |
| 28 | Edinburgh (The University of), The Faculty of Medicine |
| 29 | Glasgow (University of), Faculty of Medicine |
| 30 | St Andrews (The University of), Faculty of Medical Sciences |
| **Wales** | |
| 31 | Cardiff University, School of Medicine |
| 32 | Swansea University, School of Medicine |
| **Northern Ireland** | |
| 33 | Queen's University Belfast, Faculty of Medicine and Health Sciences |
| | |

Of the 420 papers reviewed only 7% took place in the UK, 85% were from North America; the remainder were from the rest of Europe, Asia and very few from Africa. This was an interesting geographic distribution as apart from North America, less than one-sixth of the publications were from the rest of the world. As the review was limited to English language, the greater number from N. America may reflect a publication bias that prevents a fuller picture of FD from an international perspective. However, this N. American dominance was noted in other reviews (Koppel et al, 2001; Freeth et al, 2002) and it was similar to the findings of Tekian and Harris (2012) regarding the worldwide distribution of masters level programmes in medical education.

The 30 published papers on medical FD in the UK included ten, which focused on FD practices in medical schools, while the others were on postgraduate education. The only UK wide survey was more than 18 years ago when Biggs et al. (1994) surveyed 28 undergraduate deans; only nine reported courses specifically designed for medical teachers. The same year, a survey of 186 teachers at a British medical school reported that most teachers supported the concept of training courses on teaching but only 5% believed their teaching ability to be below average (Finucane et al, 1994). Two examples of published FD interventions in the UK are the Teaching Improvement Project (TIPS) at the University of Nottingham medical school (Dennick, 1998) and the leadership capacity in medical education at the School of Medicine, University of Southampton (Hill & Stephens, 2005). Overall, published information on FD activities for medical school teachers in the UK was sparse.

In contrast, across N. America there are several publications describing FD activities at various medical schools, such as the Medical Education Scholars Program (MESP), Master Teacher Fellowship Program (MTFP) and the Teaching Scholars Programs (TSP), to mention a few (Cooke et al, 2003; Gruppen et al, 2003; Muller & Irby, 2006; Searle et al, 2006; Wilkerson et al, 2006). One could argue that the USA has more medical schools than the UK as it has 133 schools (four times the UK number) that award the Doctor of Medicine (MD) degree. However, this cannot be the only explanation as Canada has only 17 medical schools and surveys of FD in Canadian medical schools over three decades (1987, 1997, 2007) have shown a positive trend with FD activities now present in all 17 medical schools with evidence of increased funding and all schools but one have a FD committee. The surveys showed a continued emphasis on teaching but with broader coverage of other faculty roles but the absence of FD impact studies still remained a notable finding (McLeod, 1987; McLeod et al, 1997; Mcleod & Steinert, 2010). More to the point is that there were more publications on FD from Canada compared with the UK meaning their educators are at least aware of what is available and on-going regarding FD. The same cannot be said for the UK, as the literature review showed limited research on the views of medical faculty on FD initiatives in the UK.

It can be argued that the paucity of published papers is not the same as lack of activity or provision, but this is conjectural. With the increased number of UK medical schools, the increasing complexity and pressures of healthcare delivery, new approaches to teaching and competing demands on teachers' time, educators require a broad range of teaching strategies that can be used in various settings.

The question is ‘Does this mean that there is less FD provision in UK medical schools?’ I explore this question further in Phase 1 of my research (Chapter 5).

## 2.8 Summary of literature review

The review of the literature has demonstrated a need for FD and the potential usefulness of activities tailored to teaching in health care settings in developing outstanding teachers and scholars. FD has evolved over the last four decades to cover a wider area of perceived faculty needs, but teaching remains an important core area. However, despite the 40 odd years of FD, there was still a major transatlantic difference. In N. America, there was a vast amount of literature on FD and a variety of available activities with demonstrable impact in some areas, while in UK medical schools there was a paucity of literature on UK FD opportunities with minimal research on the impact of FD.

FD activities represent complex interventions in complex settings where many intervening, mediating variables (e.g. personal attributes, teacher’s status / responsibilities) interact with uncontrollable, extraneous factors (Drescher et al, 2004; Steinert et al, 2006). This is one of the many reasons why evaluation of FD effectiveness is difficult (even if changes are noted, they may not definitively be attributable to the programme); hence new research methodologies are required. Furthermore, it is also apparent that there is, as yet, little evidence on how FD is received and I believe that ascertaining the views of educators is an important step in assessing the scope for quality improvement in FD. My realist research question stated in Chapter 1 (and repeated below) will seek to address these issues.

***Does evidence provided by the stakeholders suggest in what circumstances (C) faculty development programmes (M) can be effective for training medical educators and lead to good educational outcomes (O)?***

With this question in mind, I will now explore how the literature review has helped in developing the realist framework by identifying possible mechanisms, contexts and outcomes.

## 2.9 Literature review and realism

As discussed in the introduction to this chapter, my analysis of the literature on FD formed part of the core design for a realist evaluation. Pawson and Tilley (1997) recommended that the realist researcher should look in the literature for evidence that will help in the selection of contexts, mechanisms, outcomes and theories. It will be remembered from Chapter 1 that contexts are the geographical and social framework of the programme under study, mechanisms come from the structure of the programme and outcomes follow from the programme. However, as the particular type of qualitative study that I am proposing has not been carried out before, and previous realist evaluation in medical education is limited, I had no template for this. In addition, most of the articles reviewed focused on the details of the FD intervention, how it was carried out and sometimes included sections on evaluation, hence, identifying contexts, mechanisms and outcomes from the interventions and the explanations for them was quite an intricate task. However, despite these limitations, I was able to identify some contexts, mechanisms and outcomes *(CMO)* from the literature review (Table 2.4) and I have given a brief description below; a more detailed description of the *CMO* is provided in Chapter 3.

**Table 2.4: Examples of Contexts, Mechanisms and Outcomes (CMO)**

| Contexts | Mechanisms | Outcomes |
|---|---|---|
| *Infra-structural*<br>*e.g. GMC and WFME initiatives*<br>The need for medical teachers to be trained<br>Standardisation of teaching<br><br>*Institutional*<br>School Policy on FD<br>Resources allocated<br>Competing demands such as research versus teaching<br>FD located in medical school or centrally<br><br>*Individual*<br>Facilitators level of expertise<br>Prior learning<br>Experiential learning<br>Time available to attend FD / Time available to practice what was learnt<br>Positive experience of FD<br><br>*Interpersonal*<br>Participatory approach<br>Shared expertise<br>Shared needs<br>Learning with peers | *Engagement*<br>Interactivity<br>Multimodal approach<br>Metacognition<br><br>*Motivation*<br>*a. Personal Oriented*<br>Career development<br>Personal development<br>*b. Teaching Oriented*<br>Personal teaching interest<br>Specific instructional skills<br>Theoretical knowledge<br><br>*Perception*<br>Perception of the value and relevance of teaching in their job and in the organisation<br><br>*Feedback*<br>On the part of the educators<br>By the FD coordinators<br>From peers<br>Individualised<br><br>*Collaboration*<br><br>*Qualification / Accreditation*<br>Recognised bodies and authorities | *Increased confidence in teaching*<br>Empowerment<br>Utilisation of teaching skills<br><br>*Improvement in instructional skills*<br>Increased knowledge of educational concepts and principles<br>Gains in skills (e.g. assessing learners' needs)<br>Promote reflection and provide feedback to learners.<br><br>*Behavioural changes*<br>Changes in teaching behaviour<br>Changes in attitudes towards teaching and FD<br><br>*Shared understanding*<br>Shared understanding of each other's role i.e. FD coordinator and educators role as teacher<br>Shared expertise<br><br>*Sense of Ownership*<br><br>*Improved credibility as educators* |

Suffice to say for now that although the *CMO* appears in three distinct columns, it should be noted that there is fluidity in where items can be categorised according to the construction of the realist hypotheses. Therefore at this point, this can only be seen as representational since there are many ways that each context, mechanism and outcome could be analysed. For example, it is worth noting that some of the outcomes could also be contexts or even mechanisms and vice versa. This is discussed further in Chapter 3.

### 2.9.1 Contexts

Context is not just the geographical location of a programme, it also refers to the spatial locations and the set of rules, norms, values and inter-relationships as well as the existing levels of knowledge and understanding which all set limits on the programme efficiency (Pawson & Tilley, 1997). Pawson (2006a) suggested four contextual layers (the four I's) that realists have to consider: Individual, Interpersonal, Institutional and Infra-structural. Examples of these from the literature are listed in Table 2.4 and discussed further in Chapter 3.

Using infra-structural as illustration, examples would include the influence of the GMC and the HEA in the UK as well the WFME worldwide on the provision of FD. Similarly, the Association of Universities in the Netherlands in 2008 affirmed that all university teachers must attain basic qualifications in teaching (Vereniging van Universiteiten, 2008). Steinert et al. (2006) also mentioned another important context; the influence of organisational characteristics, as well as the impact of FD on the organisation. Furthermore, there is suggestion that contextual factors are mandating a greater focus on FD as discussed at the 1st International Conference on

Faculty Development in the Health Professions held in Toronto in May 2011, with over 300 participants from 28 countries. The importance of context in education and learning was again emphasised at the 2nd International Conference on Faculty Development in the Health Professions held in August 2013, in Prague, Czech Republic.

### 2.9.2 Mechanisms

Mechanisms describe what it is about programmes and interventions that bring about any effect. The process of how subjects interpret and act upon the intervention stratagem is known as the programme mechanism and it is the pivot around which realist research revolves (Pawson & Tilley, 2008). The concept of mechanism is developed further in Chapter 3; however, some examples of mechanisms such as motivation, perception, engagement, feedback and collaboration derived from the literature are listed in Table 2.4. I describe some studies from the literature to illustrate some of these mechanisms.

Harden & Laidlaw (2012) suggested that a key factor considered to be at the heart of good teaching is passion: passion for the subject and passion for sharing the learning with students. Studies of the skills and attributes of excellent teachers highlighted passion and enthusiasm as key attributes. In a study of medical students' perceptions of what makes an effective medical teacher, the two highest-ranking attributes reported by students were 'passionate about teaching' and 'motivates and inspires the students' (Kua et al, 2006). Similarly, a review of exemplary university teachers found that they enjoyed teaching, showed enthusiasm for their subject, and

made an earnest attempt to promote students' learning (Hativa et al, 2001). Harden and Laidlaw (2012) clearly agreed as they stated,

> ***"Passion in teaching is not a luxury or a frill that we can do without; it is the key element in students' learning. When the quality of students' learning is compared in different situations, the differentiating factor is frequently found to be the passion of the teacher – more than their knowledge of subject matter, more than the teaching strategy adopted and more than the learning technology adopted." (Harden and Laidlaw, 2012, p. 20)***

This can be looked at in various ways: does FD deliver its courses with passion (i.e. a mechanism) or does it facilitate the development of passion in participants (i.e. an outcome)? Moreover, is it passion or engagement? Or is passion one facet of engagement? I explored all this further under mechanisms in Chapter 3. Suffice to say at this point that passion (and / or engagement) is a possible mechanism to be considered in more detail under the hypotheses of the inquiry (Chapter 3).

Baroffio et al. (1999) studied 88 teachers, all of whom had attended a Level I workshop and 44 attended a more advanced Level II workshop. The Level I workshop was a three-phase preparation for teaching that involved experiential and interactive learning; the Level II workshop was optional and addressed difficult tutorial experiences. Teachers who attended the voluntary Level II workshop had higher baseline scores than the group attending Level I, suggesting that these teachers (with higher baseline scores) were more motivated to improve. This prompted Steinert and colleagues (2006) to suggest that the issue of motivation to attend FD activities remains an unanswered question, they asked: What motivates participation in FD? What determines whether someone will take advantage of specific interventions at a particular time?

Other studies specifically examined the use of feedback as an intervention strategy and found that systematic and constructive feedback resulted in improved teaching performance. Marvel (1991) provided individualised feedback, based on a 45-minute videotape to each faculty member following baseline data collection and found teaching behaviour improved. Skeff (1983) also showed that feedback improved teaching performance. Litzelman et al. (1998) study highlighted the complex interactions that may occur between the FD intervention, teachers' experience and perceptions of their teaching. Steinert et al. (2006) affirmed that perception is important in FD as faculty members need to practice what they learn, and that immediate relevance and practicality is key to a good outcome. Steinert and colleagues (2006) also commented on the importance of the multimodal approach in FD interventions. All the FD interventions they reviewed used a wide range of instructional methods (e.g. small group discussions; interactive exercises; role plays and simulations) because of the need for different methods required to meet diverse objectives as well as to accommodate different learning styles such as the deep/surface/strategic approaches described by Entwistle et al (2001) and the activist/reflector/theorist/pragmatist styles described by Honey and Munford (1992). Taken together, the studies mentioned above demonstrate possible causal mechanisms for the FD outcomes reported even though they were not couched in realist terminologies.

### 2.9.3 Outcomes

Outcomes follow from mechanisms acting in contexts and provide the key evidence for the realist evaluator to mount, monitor, or modify a programme (Pawson & Tilley,

1997). Examples of outcomes reported in the literature are listed in Table 2.4 and discussed further in Chapter 3.

## 2.10 Summary

This chapter has set the framework for the realist evaluation and shown that the history of FD is one of both challenge and opportunity. Since its inception, FD has shown its capacity to anticipate and respond to changes and to sometimes act as a lever of change. It has evolved from individual to collective development, from singular to multidimensional purposes and from largely uncoordinated activities to more defined provisions. The extensive literature on evolution and content of FD described the nature of and explained the need for FD in medical schools while in the evaluation section I reviewed the articles from the view of a realist researcher. The literature revealed that there is little consistency in the way that programmes are offered; the skills and experience acquired by participating faculty differs from programme to programme and attendance is not mandatory in majority of cases. As this was an innovative approach, the literature review process was a tentative attempt to seek out evidence that was relevant to FD, while at the same time looking for material that could form the structure of a realist evaluation. From the literature, it has been possible to distil contexts, mechanisms, and outcomes, which can be used in developing the framework of a realist evaluation of FD. However, before this can happen in detail, it is necessary to explore the reasons for the choice of a realist evaluation and this is the subject of the next chapter.

# CHAPTER 3
# THEORETICAL FRAMEWORK

THE RATIONALE FOR USING REALIST EVALUATION

## 3 Introduction

The literature review described in the previous chapter has provided the background and existing evidence for the research study, which is about evaluating FD activities for training medical educators how to teach medical students. The next task is to identify an appropriate theoretical framework and the principles of realist research, as described by Pawson and Tilley (1997), appeared to align with the aims of this study. However, the use of realist research for evaluation in medical education is not common and therefore needs careful consideration. First, I need a framework with sound scientific methodology to ensure that the evaluation will be authentic. Second, the research will be applied research since the ultimate aim of FD is to improve the training of medical doctors and therefore patient care.

This chapter begins with a detailed reflection on the underlying philosophy of realism in terms of its ontological (nature of reality) and epistemological (knowledge of the world) stance. My purpose is to justify the use of realism in my research by comparing it with other philosophies. I will discuss the principles of realism followed by an analysis of the nature of causation (which is central to realist research) in an attempt to justify its use in the evaluation of FD. This will be followed by a discussion of the model that will be adapted for use in the study. The final section describes the design of the research study and a critique of the realist philosophy.

## 3.1 Realism

Realism is defined by Phillips (1987, p. 205) as “the view that entities exist independently of being perceived, or independently of our theories about them.” Maxwell (2012) expanded on this definition by explaining that realism combines the ontological view that there is a real world that exists independently of our beliefs and constructions with an epistemological stance that our knowledge of this world is our own construction, created from a specific vantage point. Therefore while our knowledge of the world is inherently a construction from a particular perspective, there is nonetheless a real world, which can be understood in both psychological and physical terms.

> ***“There is no possibility of our achieving a purely ‘objective’ account that is independent of particular perspectives. All knowledge is thus ‘theory-laden’, but this does not contradict the existence of a real world to which this knowledge refers.” (Maxwell, 2012, p. vii)***

Realism provides for social and educational science an alternative to philosophical positions such as positivism or interpretivism, which have been found wanting sometimes in explaining causation in complex open systems (Sayer, 2000). Realism challenges both the law-finding natural science methodology of positivism and the subjectivism of the interpretive approach i.e. it is neither nomothetic (law-seeking) nor idiographic (concerned with documenting the unique) and thus provides a third way between the polar ends of positivism and interpretivism (Sayer, 2000). In addition, realism is able to embrace the concepts of complexity, dynamics and change in social programmes by accepting that there is no ‘absolute’ truth. As Pawson (2013) stated,

*"The end result will be partial knowledge about partial improvements we can make in the delivery and targeting of social interventions." (Pawson, 2013, p. 112)*

### 3.1.1 Realism and FD evaluation

Before I provide a detailed examination of realism, it is important to take into account certain factors that are particular to this study on FD. The first factor is the form of realism that will be discussed since, according to Scott (2000), various interpretations of realism have been described including 'critical realism' (Campbell, 1988; Maxwell, 2012), 'experiential realism' (Lakoff, 1987), 'subtle realism' (Hammersley, 1992) and 'natural realism' (Putnam, 1999). For the purposes of this study the variant 'critical realism' is probably the most apposite since it questions the social practices that it studies (Maxwell, 2012). Roy Bhaskar (1975; 2002), the influential realist philosopher, used this term (critical realism) and argued that while there might be overabundance of explanatory possibilities, social science has to be critical of false explanations. This study of FD and its evaluation can be considered a critical review of an educational theory as it provides FD coordinators (and by extension the medical school) the rationale to effect any necessary change. However, like Robson (2011), I have chosen to just use the term 'realism' for reasons of simplicity.

The second factor peculiar to this study is that while realism as a form of scientific explanation has a long history, realist evaluation in medical education is relatively new. There have been some studies in medicine using realist evaluation. The realist framework was used to explore a cardiac rehabilitation program (Clark et al, 2005), Byng et al. (2005) used it to evaluate interventions for patients with long-term mental

illness and Greenhalgh et al. (2009) used realist evaluation to understand whole-scale transformation of health service. Other authors (Tolson et al, 2007; Redfern et al, 2003) also chose this method because it focuses on developing an explanatory model that encompasses the mechanisms of action as well as the context in which these mechanisms are activated.

However, I have been able to find little evidence of realist evaluation being used to any great extent in medical education. Ogrinc and Batalden (2009) used a realist approach to evaluate the teaching of internal medicine resident physicians while Hollenberg et al. (2009) applied the realist principle to interprofessional education within hospital settings. Matthews (2003) suggested that realist methods could be used to evaluate the work of educational psychologists and Timmins and Miller (2007) discussed how realist evaluation might be used to assess innovative practice in education. However, not all the studies above were truly realist in their methodological description. For example, Ogrinc and Batalden's (2009) gave some plausible mechanisms and contexts as well as described the outcomes but there was lack of detail about the data collected, their analysis and no CMO theory generated. Hollenberg et al's (2009) while more detailed in their analytic description lacked reference to a specific theoretical perspective, which is important, as realist evaluation is a theory-driven inquiry. So despite identifying some contexts and outcomes, they failed to explore mechanisms and ultimately failed to generate a CMO theory. In essence, there is a dearth of educational research that has explicitly used the logic of enquiry based in the philosophy of realism.

Similarly, Pawson (2003a) concluded that the use of the realist approach in social sciences seems to have been largely confined to the field of crime prevention: Pawson (2002a) and Tilley (1993) are two examples. As noted by Timmins and Miller (2007), social programmes that are evaluated in crime prevention (e.g. placing CCTV in car parks: Tilley, 1993) are less complex than programmes in education. Therefore, since realist evaluation in medical education is a relatively new venture, it is necessary to design and develop an approach within the realist paradigm that is suitable for educational research. A similar call has been made by Wong et al. (2012) to apply realism in medical education.

In the following section, I am going to discuss the key features of realism and how they apply to FD. Where relevant, I will compare and contrast these features across the three paradigms of realism, positivism and interpretivism explaining the differences and discussing similarities. I use the term 'interpretivism' as a generic term as described by Cohen et al. (2011) to include different approaches (e.g. phenomenology, symbolic interactionism and constructivism) that reject the positivist belief that human behaviour is governed by general universal laws and, instead, hold that the social world can only be understood from the viewpoint of individuals who are part of the on-going action being investigated.

### 3.1.2 Ontology and epistemology in realism

Ontology is the nature of reality and whereas the positivist holds that the world exists in an objective form, for the interpretivist the world exists only as people interpret it (Crotty, 1998). Thus for the positivist, knowledge of the world (epistemology), exists independently of thoughts but the interpretivist believes that the world exists only

through people's understanding of it. Realists, however, make a distinction between the intransitive and transitive dimensions of knowledge (Sayer, 2000). The objects of science form the intransitive dimension of knowledge while the theories and discovery of science are part of its transitive dimension (Sayer, 2000). The world itself, the intransitive dimension, remains the same even though theories about it may change. For example, when scientists changed their view (the transitive dimension) and decided that the sun was at the centre of the planetary system and not the earth, then the nature of the solar system (the intransitive dimension) did not change, only our understanding of it. Scott (2000) describes the relationship between our understanding and the intransitive dimension as changing knowledge of unchanging objects.

This is different from the principles of interpretivism, which is that there are no objective independent variables, and of positivism, that what we see is the world as it is (Crotty, 1998). Both philosophical traditions, in different ways, conflate ontology and epistemology. For empiricists, people's senses are cleansed of any preconceptions: what is presented to them is the world as it is. This means that people's knowing of the world is the same as what it is (i.e. what has been presented to the senses) and epistemology (how people gain knowledge / understand the world) is to a large extent bypassed. For the interpretivist, in contrast, epistemology is all there is since there is no distinction between thought and reality (Scott, 2000).

The realist's distinction between the transitive and intransitive dimension is easier to comprehend in the planetary system but less easy to understand in a complex social system such as FD. Social systems, as noted by Sayer (2000), are constructed by

people and therefore cannot be said to exist independently of at least some knowledge and contribution from the people. However the realist view of knowledge is that knowledge is always partial, incomplete, and fallible. This applies not only to our knowledge of the 'external world', but also to our knowledge of our *own* minds. We never have a complete, objective understanding even of our own thinking, and often discover that we were unaware of, or mistaken about, some of our own beliefs, values, and intentions (Maxwell, 2012). Sayer (2000) also remarked that when researchers change their minds, this is unlikely to produce significant changes in the phenomena they were studying. Hence, although there is some interplay of ideas between the researcher and the object of their research, this does not mean that the subject-object distinctions collapse.

The realist's interpretation of the social world is important for me since one of the purposes of this research is that the result can be used by FD developers to inform their design and delivery of FD activities. A positivist approach would have been unsuitable for me because, I subscribe to the view of Bunniss & Kelly (2010) and Davies (2007), that the reality of FD is constructed inter-subjectively through the meanings and understandings developed socially and experientially in specific contexts. This is important, as there is no ultimate way of knowing the exact impact of FD on educators' learning because of the difficulty in completely separating FD from other learning experiences of educators, as well as educators themselves having various interpretations of FD. Hence positivism, which fails to take account of people's ability to interpret and construct their experiences and their world, would not be appropriate.

Moreover, because of the epistemology of interpretivism, research based in its approach will largely result in knowledge that is personal, subjective and unique. This is because knowledge generation happens when relevant insights emerge through researcher-participant discourse (Cohen et al, 2011; Coffey & Atkinson, 1996). While this dialectical process may lead to a more sophisticated understanding of the social world, it has a narrow micro-sociological perspective i.e. it neglects the wider social contexts and constraints. In contrast, realist research, by placing the social programmes in the intransitive dimension, is able to arrive at findings that can have validity in other settings.

### 3.1.3 Stratification

A second tenet of realism, and one, which accommodates the complexity of social programmes, is that ontology is stratified. Bhaskar (1989) distinguished between three ontological domains: actual (action in events), empirical (sensory experience), and real (causal powers separate from but not always evident in empirical and actual). Summarising the work of Bhaskar, Elder-Vass (2010) noted that,

> ***"The empirical domain includes those events that we actually observe or experience and the actual is the domain of material existence, comprising things and the events they undergo. The real also includes 'structures and mechanisms' that generate those events." (Elder-Vass, 2010, p. 44)***

Realists see objects (whether they are physical, like minerals, or complex social systems) as characterised by structures and powers at different levels, which can either be activated or remain dormant. In the evaluation of FD, stratification can be demonstrated through one of its basic components, that of a teaching session. There

are a number of elements involved in a teaching session, including the level of experience of the tutor or facilitator, the amount of time he/she spends with the participants, the physical quality of the working environment, and the way that these elements combine will materially affect the impact of the teaching session. According to Lawson (1997), social systems are open and contingent. They commonly involve dependencies or combinations, which affect the elements (aspects), and the form and structure of the elements causally influence each other and therefore the whole system. Stratification allows for understanding at different levels and hence contributes to the unravelling of complex programmes.

### 3.1.4 Emergence

A third tenet of realism is the believe that the world is characterised by ‘emergence’ (Sayer, 2000), i.e. when features, aspects or objects combine together, this gives rise to new concepts that have properties which are irreducible to those of their constituents. An example from the physical world is the emergent properties of water, which are quite different from those of its constituent parts, hydrogen and oxygen. In the same way, social concepts (speech, for example) emerge from biological and physical strata but conversation cannot be reduced to its physiological processes. What the phenomenon of emergence tells us is that we have imperfect control over the outcomes as programmes not only work to change behaviour but they may also change the conditions that made the programme work in the first place (Pawson, 2013). FD activities are complex web of structures and elements and the nature of the FD intervention varies markedly as these change and combine; it is this that makes its evaluation a challenge. However, stratification and emergence

provides a strategy that enables the realist researcher to acknowledge this complexity within the research design.

### 3.1.5 Causation in realism

A fourth tenet, and arguably realism's most distinctive feature is to look for causal powers within the objects, agents or structures under investigation (Pawson, 2006a). This analysis of causation is, for Pawson and Tilley (1997), the key differentiating point in realism as it rejects the standard Humean successionist view of regularities among sequences of events (Hume, 1978; Sayer, 2000). Realism attempts to show that the usage of such explanatory strategies can lead to a progressive body of scientific knowledge. This contrasts with the stance of the positivist for whom causation involves identifying a model of a regular succession of events and seeking putative social laws. The positivist researcher seeks to gather data on regularities and repeated occurrences and to conclude that, because B follows A, then A causes B. Mohr (1982) labelled this approach to causal explanation 'variance theory'; it defines causality as a consistent relationship between variables. However in education, it is highly unlikely that indisputable causality is ever completely discoverable and causality is probabilistic rather than deterministic (Cohen et al, 2011). However, I am aware that other authors McManus (2003), McManus et al. (2004) while avoiding claims to complete discovery, have thoughtfully applied path analysis or causal modelling to medical education and their probabilistic work is amongst the best that positivism has to offer. I discuss path analysis later in section 3.1.7.

For realists, the explanation of causation depends not on the number of times we have observed regularity, but on identifying causal mechanisms and how they work and in what conditions they are activated. This was summarised by Sayer (2000),

> ***"Consequently for realists, causation is not understood on the model of the regular succession of events, and hence does not depend on finding them or searching for putative social laws. The conventional impulse to prove causation by gathering data on regularities, repeated occurrences, is therefore misguided; at best these might suggest where to look for candidates for causal mechanisms. What causes something to happen has nothing to do with the number of times we observe it happening. Explanation depends instead on identifying causal mechanisms and how they work, and discovering if they have been activated and under what conditions." (Sayer, 2000, p. 14)***

According to Cohen et al. (2011), the inferential, conjectural and probabilistic nature of much causation in educational research (rather than being absolute, deductive and deterministic) renders the study of causation problematic for educational researchers. Realists try to resolve this by using explanatory (or generative) causation, which can be illustrated through an example of a study of crime prevention cited in Pawson and Tilley (1997). The study showed that if, in a neighbourhood where there was a high rate of crime, valuable goods were marked with the householder's postcode, then a reduction in crime rates followed. The positivist would say that property marking leads to a reduction in crime rates. However, research does require going beyond the outcomes alone in order to have a greater understanding of why the crime reduction occurred i.e. what factors led to the reduction. Therefore, realists wanted to know the reasons for the relationship between property marking and crime reduction and constructed realist theories, which might explain this. Theories included: thieves were deterred because if marked goods were recovered thieves could be identified, property marking deter

burglars because of the anticipated difficulty of disposing of the goods, and the increased police presence in the neighbourhood, as they did the property marking, might be a restraint on thieves.

In realist research, the views of the stakeholders are seen as critical since they participated in the programme and hence had an understanding of it. Interviews with residents and burglars (stakeholders) found that most burglars paid little attention to the disposal issue of marked goods and the study concluded that the reason for the reduction in crime rates had more to do with the increased police presence in the area (with the accompanying publicity) than to do with the actual property marking. In an evaluation of a complex social or educational programme, the realist perspective on causation helps to identify how the programme works (or fails). Little (2010) provided a particularly trenchant statement,

> ***"It is very important to arrive at deeper understandings of the metaphysics of social causation. This means, first, understanding the complete inadequacy of the traditional positivist interpretation of causation: 'causation is no more than regularity.' This Humean view does not serve the natural sciences well, and it certainly does not help us when it comes to social causation. So it is necessary to explore a different model of causation that fits better with what we know about the actual workings of social processes." (Little, 2010, p. 218)***

Realist research informed by explanatory causation is able to avoid some of the problems of using positivist designs with complex educational programmes. The positivist researcher seeks ideally to identify an unchanging causal power and constant external conditions in which the causal power operates (Scott, 2000). Although this may be achieved in the closed system of laboratory conditions, outside in the more open situation of complex social programmes, neither the causal power

nor the external conditions are constant. FD involves individual human behaviour, relationships between individuals, the structural properties of the education system, and the organisations involved, all of which can change across time. For example, FD facilitators have different individual skills / expertise, the relationship between FD coordinators and educators may change and changes in resources (e.g. at institutional or national level) may lead to changes in programme structures. In addition, the external influences on FD such as the government, the medical education system, and the structure of the health service can vary as well as the leadership and values of the institution. So the observations / reports in the literature that some FD programmes lead to good outcomes for medical education (e.g. as reported using Kirkpatrick's levels in Chapter 2) might not be what it seems. It might be that only certain aspects of FD such as the level of competence of the facilitator or the type of educators attending contributed to good teaching outcomes or the outcomes might be facilitated by the context (e.g. the setup of the programme within the organisation) in which it operated.

In open systems, because two events occur in conjunction, we cannot always suppose that object X causes event Y unless we examine what aspects of X work and do not work and in what conditions X may or may not operate. Mohr (1982) called this approach to causal explanation 'process theory' in distinction from 'variance theory' described earlier. Process theory deals with events and the processes that connect them; it is based on an analysis of the causal processes by which some events influence others. Such a process approach to causation has been advocated by other qualitative researchers (Erickson, 1986; Weiss, 1994). Weiss (1994) argued that,

> *"Quantitative studies support an assertion of causation by showing a correlation between an earlier event and a subsequent event. An analysis of data collected in a large scale sample survey might, for example, show that there is a correlation between the level of the wife's education and the presence of companionable marriage. In qualitative studies we would look for a process through which the wife's education or factors associated with her education express themselves in marital interaction." (Weiss, 1994, p. 179)*

The realist framework seems more appropriate for evaluating a complex FD programme since a realist inquiry seeks not to establish regularities, but to explain connections in terms of the conditions in which causal mechanisms produce outcomes. The relationship between causal mechanisms and their effects is not fixed, but contingent; it depends on the context within which the mechanism operates (Pawson & Tilley, 1997; Cartwright, 1999).

### 3.1.6 Meaning in social phenomena

The fifth tenet of realism is understanding meaning. Meaning is commonly understood to be a mental rather than a physical phenomenon ('meaning' and 'mind' both derive from the same Indo-European root); in linguistics, 'meaning' is what is intended by the speaker or understood by the hearer (Maxwell, 2012). In addition, meanings are related to material circumstances and practical contexts in which communication takes place and to which reference is made (Sayer, 2000). Sayer (1992) argued that,

> *"Social phenomena are concept-dependent….what the practices, institutions, rules, roles or relationships are depends on what they mean in society to its members." (Sayer, 1992, p. 30)*

Realists acknowledge that social phenomena are intrinsically meaningful and causal powers can be material but can also be located in reasons and intentions. For

example, placing a cross next to a candidate's name on a ballot paper is likely to involve the voter's reasoning and beliefs about political parties. Understanding the meaning behind an action is, therefore, part of realism. However, the realist shares with the interpretivist the notion that it is very difficult to measure or count meaning; instead the realist researcher attempts to understand (or interpret) the meaning behind actions. Furthermore realism shares with interpretivism the view that meaning is intrinsic to a social programme, is context dependent and should be accessed as part of the research process. However, unlike interpretivists, realists argue that this does not rule out causal explanation since meanings themselves can be causes (Sayer, 2000; Maxwell, 2012). Bhaskar (1989) stated that for realists, the mental state that mediates between the situational precursors of action and the action itself are among the causes of action. He argued further,

> ***"Unless a reason could function as a cause there would be no sense in a person evaluating (or appraising) different beliefs in order to decide how to act. For either a reason will make a difference to his/her behaviour or it will not. In the former case, it counts as a cause. In the latter case, it is logically redundant, and deliberation, ratiocination (and indeed thought generally) become practically otiose." (Bhaskar, 1989, p. 92).***

From a realist point, not only are both individuals' perspectives and their situations (physical contexts) real phenomena, they are *separate* phenomena that causally interact with one another (Maxwell, 2012). Manicas (2009) stated that,

> ***"At the heart of social science explanation is the idea of a social mechanism with persons as causal agents…since persons are the dominant causes of what occurs in society, the first problem for the social scientist is to understand action as it is understood by the actors." (Manicas, 2009, p. 30)***

This contrasts with the positivists who observe behaviour and sometimes do not question the intentions of the participants. In this view of science, human beings are 'subjects' and they are studied through the observation of their behaviour. Behaviourists, for example, use the methods of natural science and seek to eliminate any references to beliefs or purposes (Scott, 2000). This type of approach can lead to erroneous conclusions as illustrated in the example of property marking and crime prevention discussed earlier (section 3.1.4). The positivist researcher would have observed the property marking and the reduction in crime rates and would not have taken into account the intentions of the police officers, the intentions of the publicity or the intentions of the residents (and burglars) and the impact of all these intentions on the intervention. Even when positivists do attempt to include intention, this is not always successful since beliefs have to be interpreted as a variable and participants may not always share the same understanding of the variable. For example, if a researcher, carrying out a census, asks participants to assign themselves to one category from a choice of white, black or mixed race, the participants may refuse to accept such a system of categorisation which they see as racist and / or they may feel that they do not fit into any of these categories. It is very difficult to reduce the intentional aspect of human behaviour to variables that can be observed and measured. Therefore, even though realists may be partly naturalist, as they can sometimes use the same methods as natural science regarding causal explanation, they also diverge from them in using '*verstehen'* or interpretive understanding (Sayer, 2000). Realists accepts that individuals' meanings have *consequences*, how individuals act is influenced by how they think about and make sense of what is going on (Maxwell, 2012). This view of intentions, beliefs, and meanings as causes

is fundamental and has been affirmed not only by realists but by many other philosophers and social scientists (Robb & Heil, 2003; Menzel, 1978).

### 3.1.7 Research methods and realism

Compared to positivism and interpretivism, realism is compatible with a wide range of research methods since the choice of method is defined by the object of study and what is to be learnt about it (Sayer, 2000). Realist research embraces methods both traditionally associated with a positivist experimental approach and those usually employed by interpretivist researchers. It avoids a cookbook prescription of methods without having a scholarly knowledge of the object of study in question (Sayer, 2000). According to Pawson and Tilley (1997), the most important factor for the researcher in selecting the method is that it fits the theory under question. They claim that a realist evaluation can use data, which are quantitative or qualitative, historical or contemporaneous, from small or large samples and so forth. The use of mixed methods in the evaluation of complex programmes is supported by non-realist researchers. Greene et al. (2001), for example, stated that complex and dynamic social phenomena can best be studied through the multiple perspectives of diverse methods, rather than through the limited lens of just one. Bryman (2006) suggested that,

> ***"In the new climate of pragmatism,…. issues to do with adequacy of particular methods for answering research questions are the crucial arbiter of which methodological approach should be adopted rather than a commitment to a paradigm and the philosophical doctrine on which it is supposedly based." (Bryman, 2006, p. 118)***

Pawson and Tilley (1997) clearly stated that, in a realist inquiry, the researchers' theory is the subject matter, and data collection is there to confirm, refute or refine that theory. Therefore the use of a particular data collection strategy does not commit the researcher to its philosophical framework; while a researcher may use empiricist methods it does not mean that all the data should be analysed within the positivist philosophy. Similarly, Blaikie (2000) noted that in positivist research, for example, some researchers use qualitative and quantitative methods yet interpret them both within the positivist paradigm. As a realist, I need to choose the right methods for the different aspects of my research study but the interpretation of the results has to be within the ontological and epistemological assumptions of realist research.

### 3.1.8 Other philosophies considered

It was important for me to consider other philosophies prior to making a final decision on realism. I considered action research, as the method seemed to suit evaluation in many ways: it is democratic, collaborative and designed to improve professional practices in many different kinds of workplaces (McNiff et al, 2003). Collaboration between researchers and those who are the focus of the research, and their participation in the process, are typically seen as central to action research especially the version known as participatory action research (McIntyre, 2008). Hence action research is popular, particularly in educational studies, and its protagonists maintain that practitioners are more likely to make better decisions and engage in more effective practices if they are active participants in educational research (Koshy, 2005; Schmuck, 2006). My research did involve a collaborative and

participatory aspect in that educators contributed to interpretation and understanding of the data and hypotheses (Chapters 1 and 4), however they were not involved in the design or conceptualisation of the study. Furthermore, as discussed by the authors mentioned above, action research tends to focus on the present, local needs of the participant group (which may or may not be generalisable to other contexts) i.e. those problems that are of immediate concern to practitioners and usually involves people theorising about their practices (Cohen et al, 2011). Moreover, action research is usually based on an interpretive framework and is about developing a theory, which is a guide to inquiry and specific action in present time in contrast to a realist theory, which will be applicable (and more generalisable) once the context activating the causal mechanism is present. Therefore as action research tends to be based in the epistemology of interpretivism resulting in knowledge that is largely personal, subjective and unique, I decided against it.

I also considered using transformative learning theory. Transformative learning is a term used in educational theory to describe a process in which learners re-evaluate past beliefs and experiences previously understood within assumptions derived from others and make their own interpretations (Taylor, 1998; Mezirow, 2000). Since first introduced by Jack Mezirow (1978), other perspectives (and expansion) of transformative learning have been developed by other authors (Boyd & Myers, 1988; Grabov, 1997). Transformative learning as a process of effecting change in people's points of view have been described as comprising various stages including participation, engagement, commitment and collaboration (Southern, 2007). Other authors have also used transformative theory to evaluate FD and viewed teaching as social practice that builds on transformational learning (Dorene & Boyd, 2012; D'Eon

et al, 2000). Transformative learning as described by Mezirow (2000), typically involves a disorienting dilemma with critical assessment / reflection and learning occurring by elaborating existing frames of reference, learning new frames of reference or transforming points of view. While these might be examples of causation I wasn't convinced that it will provide a detailed picture of my research question on FD as to the underlying mechanisms and contexts resulting in the change. Furthermore, my previous experience in postgraduate FD taught me that most participants did not undergo an epochal transformative change; rather it was more of a modification of their existing teaching and I need to understand what in FD brought about this change.

Lastly, I considered path analysis as the literature does concede that some methods are better suited to the process theory of causation (Maxwell, 2012). Path analysis is a recognised positivist approach in medical education for attempting to derive explanations of complex phenomena (McManus, 2003; McManus et al, 2004). Originally developed by geneticist Sewall Wright (1921) to examine the effects of hypothesised models in phylogenetic studies, path analysis is a statistical technique used primarily to examine the comparative strength of direct and indirect relationships among variables. It is an extension of multiple regression analysis focusing on causality to determine whether or not a multivariate set of data fits well with a particular (a priori) causal model (Lleras, 2005). It helps researchers using quantitative (correlational) data disentangle the various (causal) processes underlying a particular outcome. From its original description, path analysis has evolved into a variety of structural equation modeling (SEM) programmes and computer packages now used in sociology, psychology and education (Lleras,

2005). What these extensions of path analysis do provide is a more intricate way of thinking about testing a research problem by decomposing correlations into different pieces for interpretation of effects (e.g. how does parental education influence children's income 40 years later? or how is stress, burnout and doctors attitudes to work determined by personality and learning style?).

Conceivably, path analysis (and SEM) could be applied to understand more about how FD works as its philosophy involves understanding complex relationships by analysing the correlations between variables over a period of time. The technique allows testing of theoretical propositions about cause and effect without manipulating variables. However compared to realist theory, the ‘causal’ in the modelling refers to an **assumption** of the model rather than a property of the output or consequence of the technique. That is, people *assume* some variables are causally related, and test propositions about them using the technique. If the propositions are supported, it does **NOT** prove that the causal assumptions are correct nor explain the direction of causal effects as this could be reciprocal. Realism however explains how a causal mechanism in a particular context leads to a specific outcome and is not an assumption of a modelling equation. In my opinion, given the complexity of educational processes, disentangling the interrelationships among variables is often a difficult task and even if the result indicated that the FD programme was associated with improved educational outcomes, this would not necessarily mean that the former caused the latter (as could be seen in the property marking study, discussed above).

Having considered all the above philosophies, I decided that the realist philosophy was the most appropriate for my study as it is much more effective in understanding the actual processes that are involved in particular situations. This qualitative approach is what Miles et al. (2014) called 'local causality'. They argued that,

> ***"Qualitative approach, with its close-up look, can identify mechanisms, going beyond sheer association. It is unrelentingly local, and deals with the complex network of events and processes in a situation". (Miles et al, 2014, p. 223)***

Sayer (2000) agreed with the above and suggested that causal explanation requires a realist approach to discover actors' reasoning and circumstances in specific contexts not in abstraction from them.

### 3.1.9 Summary of realist philosophy

So far in this chapter, I have attempted to show that the principles underlying realism combine aspects of the philosophy of both positivism and interpretivism and thus seem to provide the basis for the most appropriate theoretical framework for FD evaluation. The justification for the use of realism is summarised in Table 3.1. This attempts to contrast various dimensions across the three paradigms. The three columns give the impression of three discrete paradigms whereas, in fact, realism incorporates aspects of positivism and interpretivism within its underlying philosophy. Hence, the whole spectrum should probably be viewed as a continuum with realism occupying the middle ground. This is neatly captured by Maxwell's (2012) statement,

> ***"The epistemology of realism implies that no position or theory can claim to be a complete, accurate representation of any phenomenon, including research itself, and that we should view every theory from both the 'believing' and 'doubting' perspectives, looking for what insights and advantages it provides, and for where its blind spots and distortions are." (Maxwell, 2012, p. ix)***

Realism sees social phenomena as existing in the objective world with open, dynamic and complex systems (Miles et al, 2014; Pawson, 2006a). Realism explores the complexity through the notions of stratification and emergence. The interpretation of causation means that the researcher has to try to understand and explain how outcomes are generated. Realists, therefore, recognise the context-dependence of social phenomena and the need to interpret meaningful actions (Sayer, 2000). Realist evaluation is considered one of the theory-driven inquiries even though the theories generated are considered as middle-range theories. Middle range theories are theories that involve abstraction but are close enough to the observed data to be incorporated in propositions that permit empirical testing (Merton, 1967; Pawson, 2006a). More than one middle-range theory may explain the influences of context on a mechanism to produce an outcome hence the *CMO* hypothesis generation phase of the study. In the next section of this chapter, I will describe how I interpreted and designed the realist approach for use in FD evaluation.

**Table 3.1: A Comparison of Conceptions of Social Reality**

| *Dimensions of comparison* | **Positivism** | **Realism** | **Interpretivism** |
|---|---|---|---|
| *Ontology* | The world exists and is knowable as it really is. Reality is static and fixed<br><br>Ontology is *flat* since what is observed is all that exists. This conflates ontology and epistemology | Reality exists independent of social actors and observers.<br><br>*Ontology* is *stratified* and the world is characterised by emergence. Avoids the 'epistemic fallacy' by not conflating ontology and epistemiology | Reality is subjective. There is no objective reality since reality can only be constructed through a conceptual system.<br><br>This conflates ontology and epistemology. |
| *Epistemology* | An empiricist epistemology which holds that knowledge stems from sense-data inputs such as our ability to observe patterns. Objective knowledge (facts) can be gained from direct experience or observation and is the only knowledge available – invisible or theoretical entities are rejected | There is a distinction between:<br>the intransitive dimension (the object of scientific enquiry) *and*<br>the transitive dimension (our understanding of that object, including the theories of science)<br>Because our understanding of the world may change this does not mean that the world itself changes | A relativist epistemology which holds that all knowledge is relative to one's location within a set of social norms. Knowledge is subjective and socially constructed as individuals habitualise, typify and categorise their observations / experiences. Bodies of knowledge are created and institutions (fixed pattern of thoughts and action e.g. school systems, sports, religion) emerge. |
| *The role of social science* | Discovering universal laws of society and human conduct within it | Inventing theories to explain the real world and testing these theories by rational criteria | Discovering how different people interpret the world in which they live |
| *Methods of Understanding* | Identifying conditions or relationships which permit the collectivity to exist. Conceiving what these conditions and relationships are. | Exploring events in real world phenomena at various depths so that they can be understood in different ways and at different levels | Interpretation of the subjective meanings which individuals place upon their action. Discovering the subjective rules for such action. |
| *Human behaviour* | Social scientist is an observer of social reality. Respondents are treated as objects, informants or producers of data | Observable human behaviour is characterised by underlying intention and choice. Understanding this is part of the research process | The importance of viewing the meaning of experience and behaviour in its full complexity is stressed |
| *Theory* | A rational edifice built by scientists to explain behaviour | Theory here refers to mechanisms postulated as being capable of producing the events observed | Sets of meanings which people use to make sense of their world and behaviour within it |
| *Research* | Experimental or quasi-experimental validation of theory.<br><br>A deductive logic is commonly adopted where pre-existing theoretical ideas and concepts are tested. | Explanation is concerned with how mechanisms produce outcome patterns and in what circumstances.<br><br>Abductive logic is commonly used. It is a process that cycle between induction and deduction as it explains patterns of data, entertains multiple hypotheses and draws inference to the best explanation. | The search for meaningful relationships and the discovery of their consequences for action.<br><br>An inductive logic is commonly adopted starting with data collection from which theoretical ideas and concepts emerge. |
| *Research Methods* | Quantitative methods<br>Abstraction of reality especially through mathematical models | Mixed methods. The researcher chooses the method(s) which best fits the investigation | Qualitative methods<br>Analysis of meaning |

*(Adapted from Cohen et al, 2011; and Robson, 2011)*

## 3.2 Model of Realist Evaluation and FD

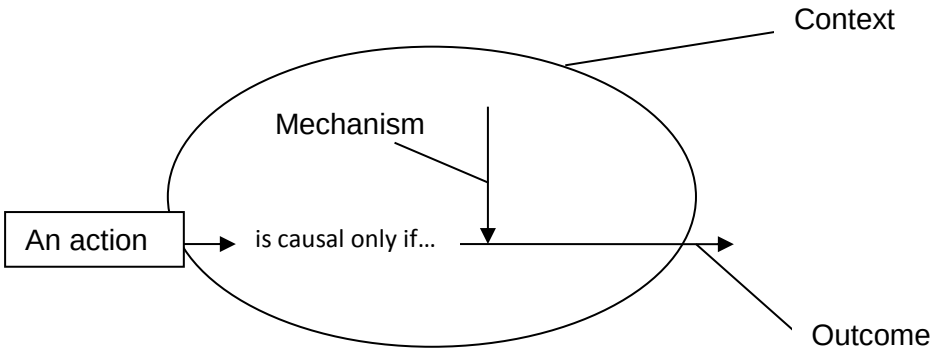
Realist evaluation has its own signature slogan ‘what works for whom in what circumstances and why’. “Why does a programme work in Wigan on a wet Wednesday but fail in Frinton on a foggy Friday?” (Pawson, 2013, p.15). This model of realist evaluation described by Pawson and Tilley’s (1997) is commonly cited as an example in public policy programmes such as crime prevention evaluation. However, the authors did not suggest that their model should be prescriptive and instead offered it as a practical approach for following the principles of realist research when carrying out an evaluation. It seems therefore, that the model may be adapted for use in educational research. This section describes the model and how it was developed for use in FD evaluation.

### 3.2.1 Pawson and Tilley Model

Pawson and Tilley gave clear directions about the principles of realist research by stating that “*causal outcomes follow mechanisms acting in contexts”* (Pawson & Tilley, 1997, p. 58). Generative causation is about outcomes being explained by the action of particular mechanisms in particular contexts (*CMO* configuration). This is represented diagrammatically (Figure 3.1) and can be explained with the example given by Pawson and Tilley (1997). If the ‘action’ in the diagram is lighting the gunpowder, then the mechanism will be the chemical composition of the gunpowder and the outcome, the explosion. However, the occurrence of the explosion will depend upon the context (i.e. the conditions have to be right – absence of damp, sufficient gunpowder, adequately compacted and oxygen present). Realist research is not about observing regularities between an action and an outcome (the lighting of

the gunpowder and an explosion) but rather about seeking an explanation for that outcome through the mechanism and the context. So, in the gunpowder example, the realist researcher would say that the heat will ignite the gunpowder (mechanism), causing an explosion (outcome) if the gunpowder is of good quality and dry (context).

**Figure 3.1: Generative causation (CMO)**




……its Outcome is triggered by Mechanism acting in Context

(Reproduced from Pawson and Tilley, 1997, p. 58, with permission)

Pawson and Tilley (1997) broadly aimed to set their model within the same basic logic of inquiry as that underpinning research in the natural sciences. Thus, theories are framed in abstract terms and are about identifying and explaining connections. Hypotheses are generated and are tested through observations. It is in the way that the theories (and hypotheses) are constituted which distinguishes a realist design because,

> ***"Theories must be framed in terms of propositions about how mechanisms are fired in contexts to produce outcomes. All else in the circumnavigation of inquiry follows from this." (Pawson and Tilley, 1997, p 85)***

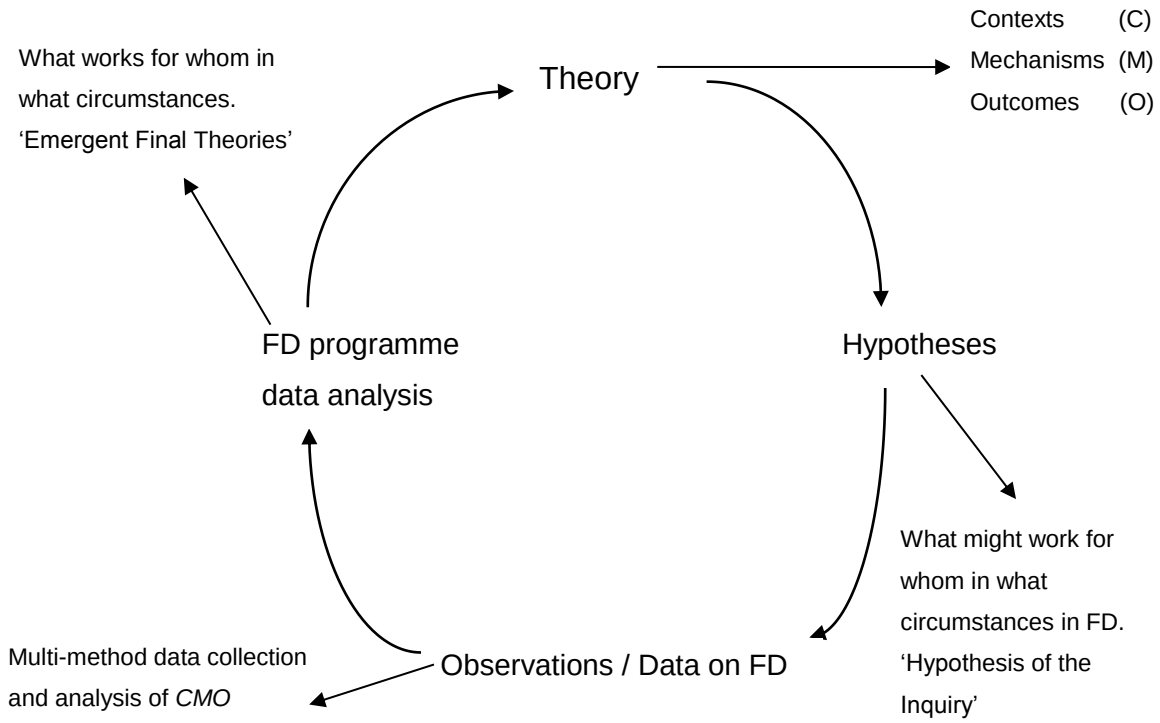
Furthermore, the way that conclusions are drawn following the testing of the hypotheses differs between positivist research and realist research. While positivist research is concerned with making generalisations and suggesting laws, the realist researcher is more circumspect in drawing conclusions and more concerned with 'specification', i.e., what works for whom in what circumstances. So, the realist evaluation of FD will not be about whether the programme as a whole 'works' but rather, which mechanisms of the FD programme, fired in certain contexts produce which outcomes.

I have adapted the realist research cycle from Pawson and Tilley (1997) as shown diagrammatically in Figure 3.2. This model can be used as a basis for the research cycle in the evaluation of FD. At the top of the research cycle is my realist theory *CMO* research question (previously stated as),

> ***Does evidence provided by the stakeholders suggests in what circumstances (C) faculty development programmes (M) can be an effective way for training medical educators and lead to good educational outcomes (O)?***

The subordinate hypotheses are the realist ideas that I developed through the phases of the study and they would be labelled as the 'Hypotheses of the Inquiry' to distinguish them from the original theory (as explained in Chapter 1). Following data collection and analysis, the *CMO* theories that emerge would be labelled as 'Final Realist Theories'.

**Figure 3.2: The realist evaluation cycle**




(Adapted from Pawson and Tilley, 1997, p. 85, with permission)

### 3.2.2 Application of Pawson / Tilley model to FD

The remaining part of this section is a description of how I used the philosophy of realism and Pawson and Tilley's (1997) model to develop a plan for the evaluation of FD initiatives. I found, as Sayer (2000) suggested, that developing a realist research plan could be a complex task as it is a multifaceted undertaking,

> ***"Unlike some of the natural sciences, we cannot isolate out components and examine them under controlled conditions. We therefore have to rely on abstraction and careful conceptualization, on attempting to abstract out the various components or influences in our heads, and only when we have done this and considered how they combine and interact can we expect to return to the concrete, many-sided object and make sense of it." (Sayer, 2000, p. 19)***

Nevertheless, I developed an outline for my research plan as follows:

First was to understand and conceptualise FD initiatives and then to abstract and define the object of study. This involved an appreciation of why FD programmes are developed and the circumstances in which they were being carried out as reported in the literature review. From this I was able to identify potential contexts, mechanisms and outcomes from which I could derive some hypotheses that might explain the connections.

Phase I of the research study involved data collection from medical school websites to understand how widespread FD programmes were across UK medical schools and to identify FD patterns and possible CMO data.

Phase II of the research study was an attempt to experience with stakeholders a FD activity with the aim of identifying connections between causal mechanisms, outcomes and the enabling (or constraining) effects of context and to explain how and why these occur. As realists also recognise that there is a need to interpret meaningful actions, understanding the intentions of the participants was integral to understanding the FD programme.

Phase III of the research study was aimed at further refining the hypotheses of the inquiry by collecting data from stakeholders (FD coordinators and medical educators) across eight medical schools and identifying *CMO* connections.

Analysis of the CMO data from all three phases led to emergent final theories on FD. Step 1 of the research plan will be discussed in section 3.2.3 below while steps 2-5 will be discussed in section 3.3 under the design plan and data collection strategy.

### 3.2.3 Understanding Mechanisms, Contexts and Outcomes in FD

According to Pawson (2006a), interventions offer resources which trigger choice mechanisms (M), which are taken up selectively according to the characteristics and circumstances of subjects (C), resulting in a varied pattern of outcome (O). Hence, in order for a researcher to explain how a programme is working or failing, the first task is to identify the mechanisms, contexts and outcomes of the programme (Box 3.1). Pawson's (2003b) in his discussion on the complexity of realist theories, advised realist researchers to focus on what they considered vital to the effectiveness of the programme. Therefore, I took great care to ensure that there were good reasons underpinning the selection of CMOs. Examples of CMOs in FD derived from the literature were listed in Table 2.4 and are discussed in further detail below before considering the hypotheses arising from the CMOs.

**Box 3.1: Contexts, Mechanisms and Outcomes (CMO)**

Contexts: Consists of the broader historical, cultural, economic, geographical, and structural factors that exist at the time of the initiative. It includes individuals (the characteristics and capacities of the stakeholders); interpersonal (relationships); institutional settings (rules, norms and customs) and infrastructural system (the wider social, economic and cultural setting).

Mechanisms: The agents of change, they describe how the structures and resources embedded in a programme influence the reasoning and behaviour of the programme subjects. So while programmes offer resources, whether they work depends on the reasoning of the subject i.e. programmes work only if people choose to make them work.

Outcomes: They represent the intended and unintended consequences of the intervention. Outcomes follow from mechanisms acting in contexts and provide the key evidence for the realist evaluator to mount, monitor, or modify a programme.

#### 3.2.3.1 Mechanisms

Mechanisms according to Pawson and Tilley (1997) are the structures of a social programme including the choices and capacities which lead to regular patterns of social behaviour. Mechanisms explain causal relations by describing the 'powers' inherent in a system, be those system substances (like gunpowder) or structures (like educational and social programmes) or agents (like policy makers). Mechanism explains what it is about the system that makes things happen (Pawson, 2006a). It is through the notion of programme mechanism that we take a step from asking whether a programme works to understanding what it is about a programme, which makes it work. Pawson and Tilley (2008) explained the role of mechanisms in simple terms,

> ***"Mechanisms describe what it is about programmes and interventions that bring about any effects. Mechanisms are often hidden, rather as the workings of a clock cannot be seen but drive the patterned movements of the hands. This realist concept tries to break the lazy linguistic habit of basing evaluation on the question of whether 'programmes work'. In fact, it is not programmes that work but the resources they offer to enable their subjects to make them work. This process of how subjects interpret and act upon the intervention stratagem is known as the programme 'mechanism' and it is the pivot around which realist research revolves" (Pawson and Tilley, 2008, p. 6).***

At a seminar, Pawson (2006b) gave the audience two questions to help to identify mechanisms.

> ***"What is it about the programme that brings about change? What resources and reasons does it offer which may influence behaviour?" (Pawson, 2006b, p. 2)***

Pawson (2006a) further explained that while programmes offer resources, whether they work depends on the choices of the subjects i.e. programmes work only if people choose to make them work.

> ***"The development of cumulative knowledge about 'what works' requires sustained investigation of the mechanism, namely the operation of choices under the inducement of programme resources." (Pawson, 2006a, p. 24)***

Programme mechanisms capture the way in which the programme's resources impinge on the stakeholders reasoning (Pawson, 2013). This approach provides a way to unravel complex issues (Marchal et al, 2012). So what is it about a FD programme that makes a difference? What goes on within FD to influence people to change? In order to identify the mechanisms of FD programmes, I used information from the literature review (Chapter 2) and information on FD on the medical schools webpages (Chapter 5). The examples of mechanisms reflect information from these sources. So using motivation as an example, a FD course may influence an educator's motivation about teaching. The educator may have originally been attending the course purely for individual gain (career, CV) or external influence (GMC directive, appraisal), but following the FD intervention, there is a change in motivation with the educator now keen to use instructional skills to help improve their students learning or the educator seeks to attend other teaching courses to improve areas of weaknesses or forms a community of practice with other educators. The mechanisms that I believed most important in FD were selected for the hypotheses of the inquiry (section 3.2.4).

#### 3.2.3.2 Contexts

Contexts as explained in Box 3.1 set limits on the programme efficiency (Pawson & Tilley, 1997). Researchers need to take into account the contexts within which programme mechanisms can be successfully fired. Realists understand that contexts operate at different levels and I considered the four contextual layers (four I's) described by Pawson (2006a):

Individual capacities of the key stakeholders: Do the educators attending the FD programme have the appropriate drive, capabilities, characteristics needed to take the intervention forward?

Interpersonal relationships supporting the intervention: How can a supportive learning environment be created? Are the lines of communication between the university and NHS management supportive or damaging to the delivery of FD by the facilitators?

Institutional setting: Does the culture, character and ethos of the organisation support FD teaching or is this overwhelmed by concerns with research or clinical workload?

Infra-structural system: The wider infra-structural system includes the political backing, government initiatives, regulatory bodies (GMC, WFME).

In order to select the contexts, I again took into account the same sources of information used above under mechanisms, as it was important to consider contexts that had been previously identified. The literature turned out to be the key source as

various authors have considered contextual factors (facilitators and barriers) affecting FD interventions (section 2.4.2). Using this information, I chose those contexts that I considered to be most important in FD for the hypotheses of the inquiry.

#### 3.2.3.3 Outcomes

Outcomes are the consequences of the intervention (Box 3.1). I viewed outcomes in terms of the aims of FD programmes and in terms of educators' needs that are being addressed as identified in the literature review (Chapter 2). Outcomes can sometimes be challenging as they can be achieved in the process of the social programme and become contexts. Thus, if FD coordinators and medical educators engage in extensive mutual exchange and communication during the FD programme, then the outcome may be a sharing of expertise between FD coordinators and medical educators (Table 2.4). However, shared expertise can also be a context. Similarly, feedback when provided by FD facilitators during the course and with participants' assessment / portfolio can be a mechanism but it can also be an outcome if educators learn effective feedback skills during an FD course and improve feedback to their own students. I explore this issue further under the critique of realist philosophy.

### 3.2.4 Developing the hypotheses of the inquiry

A basic feature of realist research is to seek substantial connections or patterns among phenomena rather than formal associations or regularities (Sayer, 2000). Pawson and Tilley (1997) did suggest that, when possible, connections need to be established first and then hypotheses should be generated to explain them but this

was not so in all the studies they used to illustrate their model of realist evaluation. In the case of FD, although there was information from the literature review that might indicate links between mechanisms and outcomes (e.g. FD leading to change in knowledge and behaviour of educators) there wasn't always specific data making such links (i.e. what exactly caused the change). The researcher, as Pawson (2002b) suggested, has to analyse the workings of the programme to discover the contexts that produce successful outcomes and those that induce failure. Therefore, to develop the hypotheses, I had to carefully analyse the data from the literature review especially studies with information on contexts and those that included evaluation data and combine these with my knowledge and experience from my involvement in postgraduate FD (section 1.2). The hypotheses were framed in terms of mechanisms activated by contexts to produce outcomes (Table 3.2). Subsequently, establishing the connections would be part of the data collection that would contribute to the assessment of the appropriateness of the hypotheses.

As noted earlier, given the presence of multiple systems and causes in the things we study and the possibility of different causes producing the same effects (the open system predicament), it is impossible to question all possible hypotheses that would occur during the period of hypotheses development. Therefore, I followed Pawson (2003b) advice to justify the choice of hypotheses by focusing on what is considered vital to the effectiveness of the programme. I choose those realist hypotheses that seem to reflect the most important aspects of FD and carefully constructed the eight hypotheses of the inquiry (Table 3.2) using information from the literature review and my personal knowledge and insight as described above.

**Table 3.2: The Hypotheses of the Inquiry**

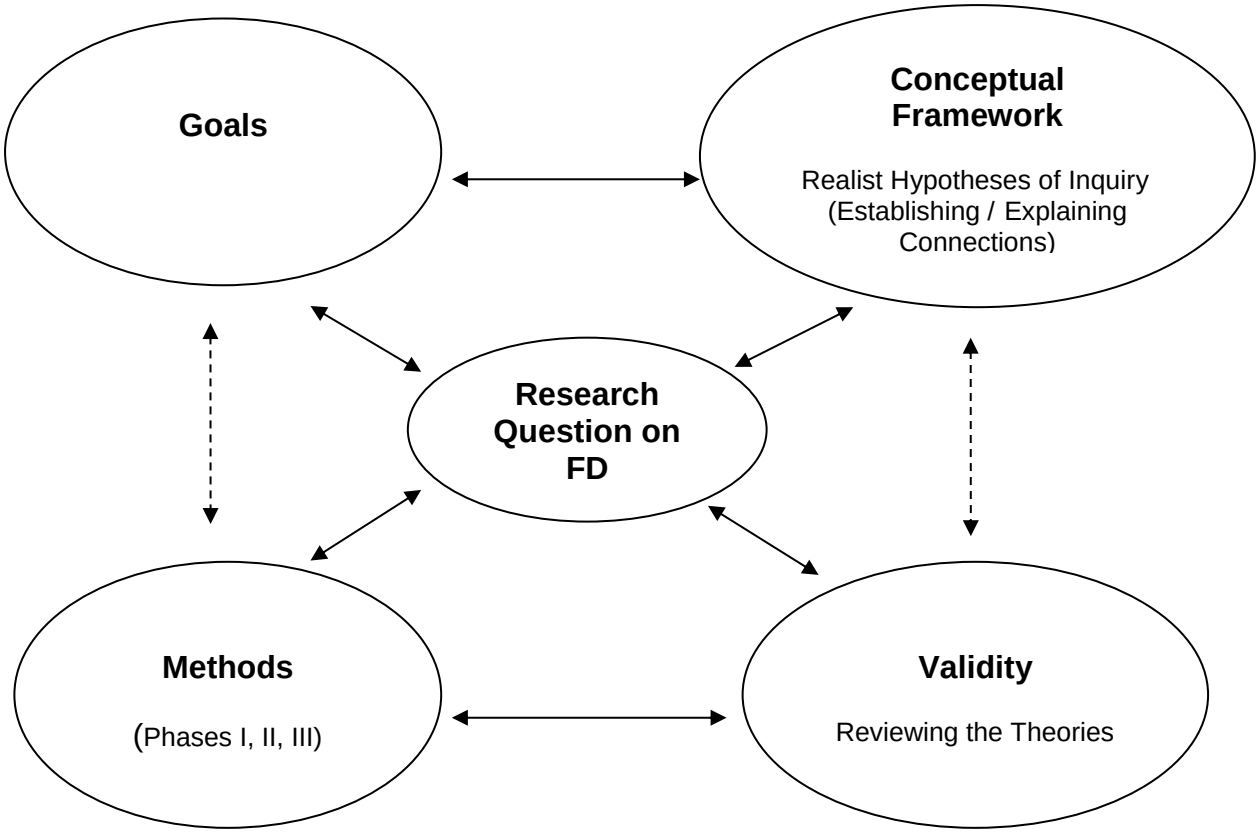
| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + | Engagement generated by the facilitator being very interactive and informative using multimodal approach | = | Increased confidence in teaching. Empowerment to utilise teaching opportunities |
| Hypothesis 2 | FD course aligned to the needs of the educator or FD relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills and change in teaching approach |
| Hypothesis 3 | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Increased enthusiasm, interest and higher importance attached to their teaching |
| Hypothesis 4 | FD using an iterative cycle of training, changes to course design and continuous dialogue with stakeholders | + | Feedback to the educators during the FD and / or assignments | = | Improved teaching performance |
| Hypothesis 5 | Sense of joint responsibility with shared needs on teaching | + | Collaboration between FD coordinators and educators in designing programmes | = | Sense of ownership, shared understanding /commitment to teaching strategy |
| Hypothesis 6 | Outside initiatives / external influences and demand to standardise medical educators teaching | + | Professionalization of teaching with qualification / accreditation and standards by recognised bodies or authorities | = | Career progression in teaching / improved recognition as educators |
| Hypothesis 7 | Time available to attend FD / Time available to practice what was learnt | + | Regular attendance and participation at FD and updates | = | Improvement in instructional skills / improved student feedback or student evaluation of teachers |
| Hypothesis 8 | FD coordinator with little or no training for the role or no time to develop the role | + | Poorly designed FD programmes on offer for educators | = | Negative outcomes for educators as no learning acquired |

## 3.3 The FD research design: an interactive model

Once the hypotheses had been selected, the next stage was to design a study that would support, invalidate or modify them. I used an interactive research model (Figure 3.3) based on the realist approach adapted from Maxwell (2013). This interactive model is an interconnected and flexible structure with each component closely tied to several others in relationships of mutual and ongoing influence (Maxwell, 2012). My research question on FD at the heart of the model, informs, connects, and is sensitive to all other components. In the upper triangle, my research

question has a clear relationship to the goals of my study and is informed by what is already known about FD (Chapter 2, literature review) and the theoretical concepts and models that can be applied (section 3.2). In turn, my decisions about relevant theories and knowledge are based on my goals and research question. Similarly, the bottom triangle is closely integrated. The methods I chose are designed to answer the research question, and to deal with validity threats to the answer. The question on FD is framed to take into account, the feasibility of the methods and the validity threats while the relevance of validity threats and the way these can be dealt with, depends on the questions and methods chosen.

**Figure 3.3: An Interactive Model of FD Research Design**




Source: Adapted from (Maxwell, 2013, p.5) with permission

The goals and research question have already been described in detail in Chapter 1, and the methods have been outlined in section 3.2.2 and discussed further in Chapter 5. Therefore the next section is about how I used the realist conceptual framework in my design to establish as well as to explain connections within the hypotheses of the inquiry. This is followed by a critique of the realist framework.

### 3.3.1 FD in a realist conceptual framework

#### 3.3.1.1 Establishing connections

The first task was to find out what was happening in FD programmes, whether the mechanisms are operating and whether outcomes occurred in conjunction with the mechanisms. For example, I needed to know whether regular feedback was happening during FD programmes and whether they were associated with improved teaching performance for example. As noted by Pawson and Tilley (1997), in order to establish connections realist researchers can use existing data, which has often been collected over a period of time. However, as such data did not exist for FD, and therefore had to be collected, the scope and the amount of data would be more limited, since the time and resources of the researcher would govern the amount of data. Nevertheless, I was keen that the collection of data should be as wide as possible since there may be other connections than those proposed in the hypotheses.

When designing a realist inquiry, Pawson and Tilley (1997) suggested that the questions that the researcher needed to ask were: who might know the data needed and how should they be asked. The authors emphasised that the stakeholders hold the information on a programme and it is the task of the researcher to identify, which

stakeholders have expertise in the area, which is being investigated. Pawson (2006a) stated that programmes (such as FD), only work through stakeholders' reasoning about the choices they make, and knowledge of that reasoning is integral to understanding their outcomes. At this stage in the research process, the information I needed was on mechanisms, contexts and outcomes in FD programmes. The stakeholders with the most knowledge on this seemed to be the FD coordinators and the educators. The FD coordinators as they have experience of how the mechanisms have worked in their medical schools and the outcomes which have followed. Educators on the other hand, have first-hand experience of attending the FD programme and are able to provide information on outcomes they deemed relevant. As Pawson (2006a) suggested,

> ***"Broadly speaking, we should expect that, in tracking the successes and failures of interventions, we will find elements of the explanation in the reasoning and reactions of different stakeholders." (Pawson, 2006a, p. 28)***

I understand it would not always be possible to collect direct information on some mechanisms, contexts and outcomes included in the hypotheses of the inquiry (e.g. hypothesis 6 resources / funding available to support initiatives) as sometimes this might not be available to the FD coordinators especially if the FD is located centrally.

#### 3.3.1.2 Explaining the connections

My intention was to collect information on the CMOs and then analyse the data in order to suggest which occurred together. The hypotheses were constructed to suggest explanations for the connections and hence there was need for data that would support, invalidate or modify them. In Chapters 6, 7 and 8, I describe the

findings from Phases I, II and III respectively, which were concerned with explaining the connections between mechanisms, contexts and outcomes to support or invalidate the hypotheses.

### 3.3.2 Review of the theories and validity

The final task, after all the data had been gathered, was to review the hypotheses and discuss the issues of validity in relation to the study. All the data were considered together and each hypothesis was examined with relevant data used to support, modify or discount the hypothesis. This will be discussed further in Chapter 9, together with the conclusions about the final theories that emerged from the hypotheses as well as a discussion of the use of realist evaluation in medical education.

## 3.4 Critique of realist philosophy

While one could say that philosophical arguments are only important to philosophers, it seems puzzling that realism has not had a more direct influence on qualitative research despite its wide acceptance in philosophy and social sciences. This might partly be due to the inherent difficulty sometimes in understanding and separating contexts from mechanisms (and even outcomes occasionally) but it might also be due to some of the ineluctable limitations of a realist approach.

First, I will start with the issue of differentiating context and mechanism. Schofield and Tolson (2010) neatly captured this when they said,

> ***"Our experience with CMO configurations has taught us that the elements of contexts and mechanisms are not always clear cut". (Schofield and Tolson, 2010, p. 4)***

Carter and New (2004, p. 14) had the same issue and using the class system as an example, they asked the question "Do we allocate class structure to context, or is context the set of contingencies which affect the saliency of the class-based motivations, together with other mechanisms acting at the same time maybe in a different direction?" For Higgs et al. (2004), class was most certainly a mechanism, since it distributed various resources and capacities, which they call 'capitals'. They argued that with the causal powers and liabilities of people as humans, these gave rise to reasons for acting in certain ways. I addressed this issue of differentiating context and mechanism by using a very tight operational definition of each item. In addition, I carried out direct observations to understand as much as possible, the various contextual factors, the interpersonal relationships and the power struggles that could make FD interventions float or sink. My view is that accurate information on the context and mechanism is important to support, modify or invalidate a hypothesis.

Second, given the challenge of complexity, the presence of multiple systems and multiple causes in the things we study, as well as the possibility of different causes producing the same effect, there is always a risk of misattribution of causality as there might be more than one mechanism operating in concert. Pawson (2013) summarised complexity with the acronym VICTORE (Volitions, Implementation, Contexts, Time, Outcomes, Rivalry and Emergence) and said,

> ***"There is no disguising the fact that the ideas driving an intervention are often multitudinous and compelling, no concealing the reality that the same intervention can trigger change in myriad ways, and no way of camouflaging the truth that the different contexts in which programmes are implemented are as wide as society is wide." (Pawson, 2013, p. 29)***

Danermark et al. (2002) further critiqued the realist mechanism concept. They argued that realists consider structure (internally related objects) and mechanism to be central concepts within the research process, however, what are the mechanisms that follow from structure, which determine the force of objects? The flaming capacity of matches or gunpowder is one thing, because this capacity is quite easy to relate to a mechanism, but using the example of organisations and paid work, they question, does the structure of paid work have the causal power of forming the life conditions of people. If this structure is about the relation of employer-employed there are by definition certain elements in these that are highly unspecific and do not say much about life conditions. The life situation of people is determined by factual conditions, such as legislations, labour market, work organisations, managers, social security system and regional conditions. Therefore, they asked is it reasonable to assume that a mechanism which reproduces the structure of paid work is triggered every time someone goes to work or applies for a job?

It seems to me that although realism acknowledges that it deals with an open and complex system, quite a substantial part of the realist framework appears to be inspired by the world of natural science (physics). Hence, when it leaves the realm of natural science and enters that of the social world, the terminology of forces and mechanism does not work quite as well in trying to explain complex matters such as organisation of paid work or FD with a teacher-learner relationship in a medical

school. According to Alvesson and Skoldberg (2009), described in these terms (forces and mechanism) social phenomena come across as mechanical and often run the risk of being overly simplified. Other authors also critiqued the ontological status of social phenomena. Whilst many would be prepared to go along with the notion of a physical and natural world independent of people's cognising experience (the law of gravity is real in the sense of being unaffected by how we choose to think about it or interpret it), phenomena in the social world seem more dependent on human meaning. Layder (1990) gave a good example of this,

> ***"Whether I chose to see a pile of stones as source of ammunition, a place of worship or as evidence of past architectural endeavour is crucial in determining the nature of my reality. Perhaps more crucially, it determines what I am likely to do – pick them up, drop to my knees or consult the guidebook." (Layder, 1990, p. 60)***

To minimise the complexity challenge, I have used an interactive, iterative, ongoing, flexible research design so that I could be responsive to the influences of the various components of the research and re-examine each hypothesis at each stage for validity. Furthermore, I am of the view that realism offers a clear alternative to research that refrains from leaving the surface level (i.e. approaches that never go outside or beyond the empirical), by analysing other aspects of a phenomenon to get to a deeper level of understanding.

Third, some critics of realism have complained that it is not always clear whether a relationship is necessary or contingent. According to Sayer (2000), this is sometimes the case, but where it occurs, it shows that we are yet to arrive at a satisfactory understanding of the situation in question, and further work on conceptualisation may be needed before we can decide whether the elements are related.

Lastly, some argue that the findings of realist evaluation are unlikely to be representative or generalisable. However, bearing in mind that the key point of realist research is local causality in context, necessary relations discovered will exist wherever their relata are present, for example, if FD is found to lead to particular outcome in certain contextual conditions facilitating a particular mechanism, then the same outcome can be expected when those contextual conditions are present. The realist approach is based on the principle that, though human agency and interaction is involved, in certain contexts, individuals are *likely*, though *not always* certain, to make similar choices about which resources they will use (Shepperd et al, 2009). In other words, particular contexts influence human choice such that semi-predictable reoccurring patterns of behaviour occur (Pawson, 2006a).

## 3.5 Summary

In using a realist evaluation I would fulfil the purposes of the researcher, which are described in section 1.4. A realist evaluation would account for the complexity of FD programmes and help to understand how some contexts can lead to successful outcomes and some can induce failure. My second purpose is to obtain results that can be transferred to other FD interventions. By using a realist approach, I will be able to identify and explain connections and understand how the mechanisms of FD interventions work and under what conditions. It should then be possible to suggest how the mechanisms of FD might work in other situations (for example FD in research or management skills) as long as the contextual conditions are present.

This chapter forms the core of the research study. Justifying the use of realist philosophy is a challenging task but a realist framework does seem to provide a tool

with which to evaluate FD programmes. Moreover, as realist evaluations do not seem to have been used extensively in medical education research, the development of the realist model for use in FD evaluation would seem innovative. It is my hope, therefore, that this would make an important contribution to knowledge by setting a pattern which others might follow. In the next chapter, I discuss the methodology underpinning the methods used for data collection.

# CHAPTER 4

# ALIGNING METHODS TO METHODOLOGY

## 4 Introduction

In the previous chapter I described why I chose to use the principles of realist research for the evaluation of FD and noted that understanding generative causation is key to realist research, which seeks to explain the connections between mechanisms and outcomes. In this chapter, I explain the methodological approach to establishing the connections between FD mechanisms, contexts and outcomes included in the hypotheses of the inquiry. Methodology is an articulated, theoretically informed approach to the production of data. It is the strategy that governs the choices behind the use of particular methods (Crotty, 1998; Silverman, 2010). My methodological approach was to explore with stakeholders, the *CMO* hypothesis; how some contexts might lead to successful outcomes and others might induce failure (Pawson, 2002b). In the first part of this chapter, I describe my sampling strategy followed by a description of the data collection strategies for each phase of the research. The research was planned in three phases: phase I was a review of FD webpages on UK medical schools websites, phase II was observation of FD sessions / interviews of FD coordinator and participants at Warwick Medical School (WMS) FD course and phase III involved interviews of FD coordinators and medical educators at eight other UK medical schools. The rationale behind the choices and the limitations of each approach are discussed in detail. In the second part of the chapter, I discuss the analytic strategy for the data collected.

## 4.1 Sampling strategy

To explain how a social programme is working, Pawson and Tilley (1997) described how researchers should use the insight and understanding of the stakeholders to help them develop the hypotheses and explanations. To be able to do this, I needed stakeholders who, collectively, would have the expertise to create a rich picture of FD interventions. However, as it was not practical to use every single stakeholder, I needed a method of purposeful sampling or selection, a strategy in which particular persons, settings or activities are deliberately selected to provide information that is particularly relevant to one's questions and goals (Palys, 2008; Maxwell, 2013). Weiss (1994) further argued that,

> ***''Many qualitative studies do not use samples at all but panels – people who are uniquely able to be informative because they are expert in an area or were privileged witnesses to an event.'' (Weiss, 1994, p. 17)***

In phase I, all 33 UK medical schools FD webpages were sampled. In phase II, I interviewed the FD coordinator (see definition in section 4.1.1 below) prior to the course. During the course, I interviewed as many educators as possible within the constraints of being a single researcher. Six months later, I purposefully selected twelve educators (six from each cohort) for in-depth interviews; purposive selection was based on the observed engagement score (section 4.2.2.1), seniority and gender. The rationale was to capture as wide a view from the educators as possible. In phase III, the purposive selection of the medical schools to interview one FD coordinator and one medical educator (from each school) was based on the FD webpage categorisation (section 4.2.1.1) result from phase I (excellent, good, average and poor webpage). I chose eight medical schools for two reasons: first this

was within the recommended range in the literature (Creswell, 2007; Britten, 1995) and second, the time constraints for a single interviewer meant limiting travelling time to three to four hours.

The stakeholders I chose were those who had knowledge of FD. The FD co-ordinators in charge of FD programmes and the educators who had attended the FD activity were the key stakeholders, as each group held different knowledge of FD. The aim, in using different groups of stakeholders, was to develop what Pawson and Tilley (1997) described as cross-fertilisation between different interpretations of the programme. This meant that different stakeholders, knowledgeable about different aspects of the CMO configuration were able to provide different but complementary views of FD. In the next section, I discuss the rationale for choosing each group of stakeholders.

### 4.1.1 The FD Coordinators (FDC)

I have used the title FD coordinators (FDC) for simplicity even though there are a variety of titles in use including director of staff development, director of medical education, clinical education facilitator, teaching and learning specialist. The FDC recruited from the eight medical schools develop, deliver, and translate the programme into practice, hence, they would have specific ideas about what works within the programme (M), experiences of successes and failures (O) and some awareness of the conditions (C) in which the programme works. They would have specific knowledge based on their own experience and the specific FD programme that they run. However, I felt they would still be able to discuss FD more widely as they have experience of managing educational programmes and were used to

considering how programmes work. In particular, they would have insight into the medical school institutional value (including mission and governance) on teaching, which is key to the provision of FD activities. As previously discussed in the literature review (Chapter 2), the FDC would have their own opinions and views of FD activities which have never been published before in the UK.

However, there possibly could be a problem with data from FDC based on their beliefs about the value of FD programmes, which might influence their views. There is the possibility of bias in their wanting to give a positive portrayal of FD activities. Nonetheless, I felt that, as long as I was aware of this issue and I was using multiple data sources, I could still obtain useful and insightful views from the FDC on the appropriateness of the hypotheses.

### 4.1.2 The Educators

When discussing the contribution of the stakeholders, Pawson and Tilley (1997) suggested that the programme participants were more likely to be sensitive to the mechanisms of the programme than to other aspects. The explanation was that since programme mechanisms provide the reasons and resources, which encourage participants to change, participants are invariably in a good position to know whether they have been so encouraged. Therefore, educators would be in a good position to know whether FD has led them to make changes.

In addition, educators who participate in an FD programme would, I thought, be able to offer a wider perspective on the FD than just a consideration of mechanisms, bearing in mind that they teach students in their own contexts. Therefore, I expected

that their knowledge would be wider than just the mechanisms and they would be able to consider other aspects (e.g. the contexts involved in implementation of teaching programmes) although it is possible that any explanations would be construed in terms of their experiences of their own teaching. Furthermore, just as with FDC, their opinions on FD in UK medical schools have never been articulated before.

However, I had to bear in mind that most educators would normally experience just one journey through a programme and, therefore, may have little understanding of all its outcome patterns, which are an aggregate of thousands of such trajectories. Nevertheless, from a realist point of view, it is the depth of understanding that is more important rather than a focus on breadth and my use of multiple data sources and the methodological and data triangulation approach that I used would help in overcoming such problem. These are as described in detail in the next section under the data collection strategy

## 4.2 Data collection strategy

I used mixed methods (glossary, pg. viii) for data collection for three key purposes. The first purpose was for *triangulation* (Maxwell, 2012). I used the different methods as a check on one another, seeing if methods with different strengths and limitations all support a single conclusion i.e. one of the hypotheses of inquiry. The second purpose was to gain information about different aspects of my research. This is what Greene (2007) called *complementarity* where different methods are used to broaden the range of aspects of phenomena that was being addressed. For example, during observation I collected both quantitative and qualitative data to describe context,

setting, behaviour and frequency of events while interview data were useful in understanding the perspectives and goals of participants. The third and probably the most important reason was to gain a greater depth of understanding. I agree with what Greene (2007) described as a *dialectic* stance for combining methods: a strategy that generates a dialogue among the results of different methods, and an engagement with differences in findings that forces one to re-examine one's understanding of what is going on. Greene argued that the use of triangulation to simply confirm a conclusion has been overemphasised and overrated in mixed method research, and that the use of different methods was most valuable for providing divergent perspectives, and thus creating a more complex understanding of the phenomena studied (Greene, 2007).

However, the majority of my data collection approaches were qualitative and there were various reasons for this. The main strength of qualitative data is their ability to elucidate local processes i.e. meanings and contextual influences in particular settings or cases (Maxwell, 2013). They are well suited for locating the meaning that people place on events and I needed to understand how participants saw the links between mechanisms and outcomes in FD. Therefore, I used observations and interviews to explore how stakeholders understood and interpreted the FD programme. Blaikie (2000) noted that qualitative methods allowed participants to develop their ideas and give their own, personalised view of the world.

Furthermore, as described by Miles et al. (2014), qualitative research is conducted in a field or life situation so the phenomenon under study is embedded in its framework. This allows the researcher to understand latent and non-obvious issues, hence the

advantage of my direct observations in phase II. Lastly, a key feature of qualitative data is their richness (Miles et al, 2014) and my aim was that the stakeholders would give rich descriptions that would allow me to gain a deeper understanding of FD. I would use the qualitative data to validate, explain and illuminate the quantitative data. In the following sections, I will describe the various data collection strategies (webpage review, observations and interviews), the rationale behind the choices and the limitations of each approach.

### 4.2.1 Webpage review

As I highlighted in Chapter 2, most of the literature on FD was from North America with very few publications from the UK. The hypotheses of the inquiry were developed in line with the literature review. So the initial question was, 'how widespread is FD provision in UK medical schools?' I realised that it was possible to gather *CMO* data on FD in UK medical schools from the webpages (where such data are available) in order to reflect a UK perspective. I therefore decided to use webpage review as an important data collection method that could contribute to the realist *CMO* framework by providing contextual and outcome data on FD in the UK. This would then be used to refine and modify the hypotheses of the inquiry before moving on to further data collection in phases II and III. This iterative, explanation building process of analysing initial data to help tailor further data collection in confirming, refuting or refining emerging hypotheses is important in realist evaluation (Wong et al, 2012).

Furthermore, the webpage is the public face of each medical school and its advertisement of the FD culture within it and the institutional value placed on

teaching is key to the provision of FD opportunities (Knight & Trowler, 2000; Healey, 2000). This should be visible and publicised on the webpages because of the increasing use of the internet in the decision-making and selection of universities by potential students (graduates and undergraduates). Similarly, non-university teachers (e.g. NHS staff) with no access to secure pages will be able to access information on relevant courses. Moreover, as highlighted in Cox and Emmott's (2007) survey, student recruitment, institutional reputation and the accessibility of information to stakeholders are the 'key drivers' of UK universities websites. Hence, there is a need for universities to have attractive and clearly understood public webpages with readily navigable information on such characteristics as programmes, courses, location and accreditations (Schimmel et al, 2010).

Webpage review is unobtrusive and non-reactive since the researcher has no influence on it, but it can be open to different interpretations hence the attempt to have an objective scoring index (section 4.2.1.1). Webpage review is useful in a multi-method strategy as it would provide another type of data which complements the data obtained from observations and interviews, which have a lot more reactivity. In this study, it would also provide some information about FD in the institution, which could be related to what an interviewee said (cross-referencing). In the next section I describe my approach to the webpage analysis using a webpage scoring index that I developed.

#### 4.2.1.1 Scoring index

My aim was to do a content analysis of the FD webpage on the website of all UK medical schools inclusive of texts, images, symbols and maps. I found Krippendorff's (2004) definition of content analysis of documents useful,

> ***"A research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (Krippendorff, 2004, p. 18)***

His definition stresses the relationship between content and context, which is important in a realist approach. The *context* includes the purpose of the document as well as the institutional, social and cultural aspects, while *content* includes not only written material (text) but also data such as works of art, images, maps, sounds, signs and symbols.

According to Robson (2011), sorting out the categories is the most crucial aspect of content analysis. Therefore, my first task was to decide on a categorising system to aid the review of the FD webpages. To do this, I developed a webpage scoring index with three broad categories and 11 criteria (Table 4.1). The 1st category *general qualities* was modified from the widely used checklist originally proposed by Alexander & Tate (1999) to evaluate information in the print media but now extended to evaluate information on the web (Alexander & Tate, 2001). The criteria under general qualities were authority, accuracy, objectivity, currency and coverage. I developed the 2nd category *web utilisation* (functionality, flexibility, navigation) and the 3rd category *FD specific* (comprehensiveness, location, contacts) myself. I developed each criterion to assess a different component of the webpage. To improve reliability I asked two colleagues to independently check on whether each

criterion accurately reflected the concept it was applied to and both individuals were in agreement.

The next task was to decide on the scoring system for the criteria. To do this each criterion had to be operationalised, that is, an explicit specification has to be made of what indicators one is looking for when making a decision whether the criterion is present or absent. There is discussion in content analysis circles about the degree of inference that may be called upon when categorising items (Maxwell, 2013). This is expressed in terms of *manifest* and *latent* contents corresponding essentially to low- and high-inference items respectively. Manifest contents are those which are physically present (e.g. a particular word or phrase); latent content is a matter of inference or interpretation on the part of the researcher (Maxwell, 2013). Bearing in mind my research question, and the importance of making the connection between content and context, I chose the low-inference (manifest) approach to achieve a more reliable result. Having decided on the manifest approach, I developed a scoring system with each criterion scored 0 (no component met), 1 (one or two components met) or 2 (all components or more than two components met) giving a maximum score of 22 points for the 11 criteria. A webpage score of 0-6 was classed as poor webpage, 7-12 average, 13-17 good and 18-22 excellent webpage (Table 4.1). This classification template was subsequently used for purposeful sampling in phase III as discussed in Chapter 5.

**Table 4.1 Webpage Scoring Index**

| GENERAL QUALITIES | WEB UTILISATION | FD SPECIFIC |
|---|---|---|
| **Authority**<br>Is it clear who is responsible for the page contents?<br>Is there a way of verifying the legitimacy of the page's sponsor /author?<br>Is it clear who wrote the material and are the author's qualifications for writing on this topic clearly stated?<br>If the material is protected by copyright, is the name of the copyright holder given? | **Functionality**<br>Are there links to other pages, sites, and do they work? Are the links consistent, do one get booted out or locked in so can't go back?<br>Does the page have local search function?<br>Interactivity: Can one jump text with buttons or links? | **Comprehensive**<br>Variety: Types of FD activities; short courses, programmes, e-learning, peer teaching, mentoring, individual.<br>Details of each activity, duration, contact time, resources, cost<br>Accreditation of programmes |
| **Accuracy**<br>Is the information reliable?<br>Are the sources of factual information clearly listed so they can be verified in another source? Is the information free of grammatical, spelling, and typographical errors? (These kinds of errors not only indicate a lack of quality control, but can actually produce inaccuracies in information).<br>Is it clear who has the ultimate responsibility for the accuracy of the content of the material? | **Flexibility**<br>Is the text readable and the typography appropriate?<br>Is there multimedia – photos, music, video?<br>Are the graphics professional, optimised, clear and crisp? | **Location of FD**<br>Departmental (within the medical school)<br>Central (within the university)<br>Both<br>External (outside the university – have to attend FD somewhere else ) |
| **Objectivity**<br>Is the information provided as a public service?<br>Does the information show a minimum of bias?<br>Is the information free of advertising?<br>If there is any advertising on the page, is it clearly differentiated from the informational content? | **Navigation**<br>Is it easy to move through the site?<br>Are descriptive subheads used to help organise information and give the reader a preview /overview?<br>Does the page load quickly? | **Contact details**<br>FD coordinator contact details<br>Name<br>E-mail address<br>Telephone<br>Course administrator contacts |
| **Currency**<br>Is the page dated? If a date is provided, it may have various meanings. It may indicate: when the page was written, first placed on the web or when the page was last revised. Are there any other indications that the material is kept current? Are the links current and do they point to existing pages? | | |
| **Coverage**<br>Intended audience: Is the information relevant to the topic? Index or table of contents or site map of depth of topic covered? How comprehensive is it – does it extensively or marginally cover the topic?<br>Is it free or is there a login or fee to obtain the information? If the page requires special software to view the information how much is missed if the software is not present? | | |

| **Score** | **CATEGORY** | |
|---|---|---|
| 0: No component met<br>1: Some evidence of the criterion (1-2 components met)<br>2: All components or > 2 components met | **Excellent** webpage<br>**Good** webpage<br>**Average** webpage<br>**Poor** webpage | 18-22<br>13-17<br>7-12<br>0-6 |

#### 4.2.1.2 Limitations of webpage analysis

There are some limitations in using webpage analysis as a research instrument. There might be limited (or no) obvious information on the webpage of some medical schools despite having FD activities, for example due to password protection. In addition, the full range of FD activities offered and information such as contact details are constantly updated on some webpages but could be very out of date on others. Furthermore, while webpage analysis could provide information on context for the hypotheses, one has to bear in mind the limitation of textual analysis inability to portray a rich understanding of context within which particular meanings emerge (Graneheim & Lundman, 2004).

### 4.2.2 Observations

Direct observation in a natural setting for a period of time is fundamental to understanding the medical school culture, the setting of FD and a valuable source of information (Safman & Sobal, 2004). Observations would enable me to gather data on the physical, human, interactional and programme settings, which would help in understanding the influence of context on mechanisms and outcomes, thereby supporting or refuting the hypotheses. The case for such qualitative observational data in the understanding of causal processes is powerful (Cohen et al, 2011), as observational data are sensitive to contexts and demonstrates strong ecological validity (Moyles, 2002).

Observation provides a direct and powerful way of learning about people's behaviour, views, attitudes and the context in which these occurs (Maxwell, 2013). Hence, observation would enable me to understand the context of FD, to be inductive, to move beyond perception-based data (e.g. opinions in interviews) and to access personal knowledge thereby providing a different perspective from the interviews as well as a check on my interview data. Observation would provide supportive data that would complement and set in perspective data obtained by interview. Although interviewing is often an efficient and valid way of understanding someone's perspective, observations would enable me to draw inferences about this perspective that I could not obtain by relying exclusively on the interview data. This is particularly important for arriving at tacit understanding as well as exploring aspects of participants' perspectives, which they may be reluctant to directly state in interviews (Robson, 2011). For example, observing how an educator responds to questions or discussions on the value or importance of FD might provide a better understanding of (or complement) the educator's actual views than what he or she said in an interview. As Montaigne observed over 400 years ago, saying is one thing; doing is another (de Montaigne 1588; 2004). I accept though that observation is still open to interpretation and context as educators can behave or say what they think is expected in that company in order to present themselves in a favourable light – the social desirability response bias (Bryman, 2012). Nevertheless, observation answers some of the criticisms levelled at qualitative research by Silverman (2006) that interviews are often the sole method used to gather data. As Dexter (1970) emphasised,

> *"No one should plan or finance an entire study in advance with the expectation of relying chiefly upon interviews for data unless the interviewers have enough relevant background to be sure that they can make sense out of interview conversations or unless there is a reasonable hope of being able to hang around or in some way observe so as to learn what it is meaningful and significant to ask." (Dexter, 1970, p. 17)*

To facilitate observation, I developed an observed engagement scale, which I describe in the following section.

#### 4.2.2.1 Observed engagement scale

I developed a multidimensional observed engagement scale by combining the three categories of engagement (behavioural, cognitive and emotional) into one scale. While the past decade has provided a wealth of research on engagement, much of the early engagement research incorporated definitions of engagement from one of three uni-dimensional categories (behavioural, cognitive, and emotional) and they were found wanting (Fredericks et al, 2004). Hence other authors have supported the shift to a more multidimensional view of engagement by combining two or three categories (behavioural, cognitive or emotional) into studies of multiple dimensions of engagement (Yonezawa et al, 2009). Glanville and Wildhagen (2007) conducted confirmatory factor analysis (CFA) on previous engagement research to test whether the indicators used were acceptable measures of the types of engagement they claimed to measure and concluded that a more multidimensional concept of engagement was more robust. In my view, the shift to a multidimensional view of engagement is a step forward in enriching the engagement concept. Based on that, I condensed the three categories into a 5-point engagement scale where 1

represented no engagement and 5 represented full engagement (Table 4.2). One could argue that the multidimensional observed engagement scale was still based on descriptors that are external and observable hence internal engagement will not be captured. While this may be so, I believe the multidimensional descriptors used throughout the course (over a three day period) would be enough to capture a participant's engagement pattern. Furthermore, during the development stage, and to improve the reliability of the scale, I asked two colleagues to independently review whether each descriptor reflected the concept it was applied to and both reviewers were in agreement. The piloting and application of the engagement scale is discussed later in Chapter 5.

**Table 4.2: Observed Engagement Descriptors and Scale**

| ***CATEGORY*** | ***DESCRIPTORS*** | |
|---|---|---|
| A. BEHAVIOURAL | Vigour | Participation |
| | Intensity | Attention |
| | Absorption | Task completion |
| B. EMOTIONAL | Dedication | Vitality |
| | Enthusiasm | Anxiety |
| | Interest | Boredom |
| C. COGNITIVE | Investment | Metacognition – Summarising |
| | Interactivity | Inquiry – Asking / Answering questions |

**5-point Observed Engagement Scale**

1. No Engagement: Inattentive and unresponsive
2. Minimal Engagement: Emerging / fleeting low level of engagement, some evidence of awareness
3. Partial Engagement: Emerging engagement but not sustained or unpredictable
4. Mostly Engaged: Engagement occur the majority of the time.
5. Full Engagement: Completely engaged

#### 4.2.2.2 Limitations of observation

I am aware of limitations in using observations as a research method apart from being time consuming in terms of the actual observation, developing the instruments (observation schedule and engagement scale) and negotiating acceptance (with FDC and participants). One such limitation is the reactivity effect, which is the extent to which the observer affects the situation under observation (Bryman, 2012). According to Robson (2011),

> ***"The question is how do we know what the behaviour would have been like if it wasn't observed? Moreover, whether one takes on a very detached or very involved role as observer, or something in between, there are related methodological and ethical problems. Virtual total detachment can come across as antisocial and itself cause reactions from those observed. To be highly involved risks compromising one's researcher role." (Robson, 2011, p. 317)***

To address the reactivity issue, I considered various strategies. First was the type of observation method to adopt. Ethnography (glossary, pg. viii) as originally described is an approach to the description and interpretation of the culture and social structure of a group (Robson, 2011). As an approach to the exploration and understanding of social settings and social processes, ethnography has its roots in anthropology, and tends to classically involve a complete immersion and participation in the group being studied (Delamont, 2002). However, ethnography has evolved and become more inclusive with various approaches now being used. Bryman (2012) described the following ethnographic approaches: participant observer (researcher status is covert), observant participant (researcher status revealed), non-participant observer and non-participant observer with interaction. The classic prolonged immersion

requires a substantial time commitment and was not suitable for observing cohorts of educators attending FD activities. I also decided against a covert approach as I was keen to interact and get the views and perspectives of educators on the FD they were attending. Therefore, I choose 'non-participant observer with interaction' which was described by Bryman (2012) as observing but not participating in the group core activities (in this case the FD learning process). Interaction with group members occurred through interviews (during breaks, before the day started and at end of the day). A non-participant approach during the core activities was achieved through techniques such as avoiding prolonged eye contact, not reinforcing attempts at interaction from the group and planning position in the environment to be 'out of the way' (Robson, 2011). Other authors have used the non-participant observer with interaction approach with good result. For example, Jon Swain (2004) in his study of friendship groups in a school used the approach to obtain good data and said,

> ***"My descriptions and interpretations…. are based on two major sources of data: firstly, my non-participant observations of the boys and girls during lessons, and around the school environment; and secondly, on a series of 104 loosely structured interviews… based on nominated friendship groups of between two and three pupils." (Swain, 2004, p.169)***

Having decided against a covert approach, my second strategy to minimise reactivity was to inform the participants well in advance of the FD course that I would be present at the course. They had adequate opportunity to clarify my role, what I would be observing so they were not surprised to see a researcher during the course.

The third strategy was habituation i.e. my repeated presence in the setting (from the start of the programme to the end), so that my presence was no longer new to the participants. Furthermore, Becker (1970) pointed out that an observer is generally much less of an influence on participants' behaviour than is the setting itself (though there are clear exceptions to this, such as situations in which illegal behaviour occurs). Maxwell (2013) agreed with this and suggested that for observation studies, reactivity is generally *not* as serious a validity threat as one is led to believe.

Lastly, I was aware of the potential danger of early preliminary analysis as there is a tendency for observation to generate lots of data fragments (Bryman, 2012). Having already developed the hypotheses, the issue here could be trying early to piece data fragments together into pre-existing hypotheses (expectancy effect) which could subsequently affect my thinking and interaction with the data hence I adopted an open-ended, inductive approach to minimise this effect (Cohen et al, 2011).

### 4.2.3 Interviews

One-to-one (face-to-face) interviews are personal encounters in which open, direct, verbal questions are used to elicit detailed responses (DiCicco-Bloom & Crabtree, 2006). Interviews have the advantage of encouraging responses to all questions; it is difficult to ignore questions when talking with an interviewer (Quine, 1985). Furthermore, the interview, as noted by Cohen et al. (2011), enables participants to discuss their interpretations of phenomena and to express how they understand situations from their point of view. Interviewing is also a valuable way of gaining a

description of action and events especially for events that took place in the past. As Weiss (1994) stated,

> ***"Interviewing gives us access to the observations of others. Through interviewing we can learn about places we have not been and could not go and about settings in which we have not lived." (Weiss, 1994, p. 1)***

As a clinician of many years standing, I have experience of talking with people, helping them to express their views openly and dealing with complex social interactions. I am used to listening, using clear questions and to facilitating frank discussion. Thus, I have the skills, noted by Robson (2011), needed for interviewing which include constructing and delivering questions, responding to cues and handling sensitive topics. The interviews would provide additional information to my observations and help to check the accuracy of my observation data.

As this is a realist study, the interviews would follow the principle of 'theorising the interview' (Pawson, 1996). This means that the realist theory will focus and prioritise the inquiry. The basis of the interview is the researcher's hypotheses and the purpose of the interview is for the interviewee to confirm, falsify or refine the hypotheses. The first stage in the interview is the *teacher-learner function* in which the researcher teaches the overall conceptual structure of the investigation to the respondent so that they are left in no doubt as to the underlying purpose of the research task. The respondent needs to be able to understand the general theoretical ground that is being explored and to have a clear idea of the concepts that the researcher wishes to discuss. The second stage in the interview is the *conceptual-refinement function*, when the respondent offers their own thinking on the

researcher's theories and is given the opportunity to clarify his/her thinking. As the interview progresses, the respondent is offered a description of the parameters of his/her thinking followed by opportunity to explain and clarify (Pawson & Tilley, 1997; Pawson, 1996).

During the interview the purposes and the agenda are shared between the interviewer and the respondent and, hopefully, this not only enables respondents to express their views with confidence, but also facilitates the enhancement of their thinking. Thus an advantage of face-to-face interviewing is being able to clarify responses including non-verbal cues, as well as the flexibility in presenting the items depending on the circumstances (Streiner & Norman, 2008). I believe non-verbal cues are important in my study as there would be differing views from educators which may or may not correlate with their institutions ethos. This will require rapport, probing and deeper exploration hence my decision against telephone interviews. While telephone interviews have the advantage of being cheaper, quicker and not susceptible to interviewer appearance / characteristics (hence possibly reduce bias), the lack of non-verbal cues can be a handicap. For example, visual cues which could suggest further exploration of a topic or the need to change the topic are absent. Furthermore, it is not possible to gather contextual information, for example, from observation of the respondent and the environment. For similar reasons, I decided against e-mail interviews.

Another option that I considered but rejected was a focus group interview. A focus group interview has several participants (in addition to the moderator / facilitator)

exploring a particular topic in depth with the emphasis on generating interaction and exchange of views within the group members (Bryman, 2012). Focus groups have different purposes from one-to-one interviews in that the group dynamics and responses become part of the data. As a derivative of the face-to-face interview, it is more resource-efficient and, because the results are polyphonic, the effect of the interviewer on the interviewees is lessened (Frey & Fontana, 1991). Whilst a focus group might allow me to get a wide range of educators' experiences of FD, the disadvantage of this approach would be the inability to delve deeply into individual's perceptions of FD. Other concerns would include a need for the facilitator to be sensitive to group dynamics, the risks that participants may feel pressurised to conform i.e. offer a 'public line' instead of a honest personal response and the issue of dominance by one or two people in the group (Cohen et al, 2011; Robson, 2011). There is the added difficulty of getting several people into one place at the same time (Streiner & Norman, 2008).

#### 4.2.3.1 Limitations of interviews

Nevertheless, there are some potential limitations in using face-to-face interviews as a research method. It is particularly susceptible to the effects of the appearance of the interviewer on the interviewee. As Cohen et al. (2011) noted an interview is not a simple exchange of pure information but a social encounter, which shares many of the features of everyday life. For example, as in everyday life, some respondents might feel uneasy while others might be more trusting: meanings that are clear to one might be relatively opaque to another. I know that such constraints of everyday

life would be part of the interaction and that, when processing the information, I need to be aware that respondents have different beliefs and a variety of understanding.

Furthermore, face-to-face interviewing is more expensive than telephone interviews because of the additional travel costs, as well as being time-consuming to set up and needing time free of interruptions (Chapple, 1999; Robson, 2011). However, as mentioned, realist interviews require rapport, probing and deeper exploration, hence my decision to do face-to-face interviews. A further criticism might be that my own positive views on the FD might have influenced the way I conducted the interviews but I ensured that, by setting out the design of the interviews clearly and making each step clear and transparent (Chapter 5), I moderated such influence.

## 4.3 Data analysis strategy

This was a mixed method research with both qualitative and quantitative data hence data analysis included both approaches. The analysis was theory-driven in order to find the *CMO* evidence to support, refute or modify the hypotheses.

### 4.3.1 Quantitative data analysis

Many research data are numerical and, while there is argument from some quarters that most concepts in education are simply not reducible to numerical analysis (Horkheimer, 1972), I believe quantitative data analysis is a powerful research tool which has its own place. According to Cohen et al. (2011),

> ***"Quantitative data analysis has no greater or lesser importance than qualitative analysis. Its use is entirely dependent on fitness for purpose. Arbitrary dismissal of numerical analysis is mere ideology or prejudice." (Cohen et al, 2011, p. 604)***

Bearing this in mind, I used various numerical analytic methods as appropriate in each phase of the research project including: descriptive method, slope diagrams construct diagrams and statistical analysis. These are described in detail in the next chapter.

### 4.3.2 Qualitative data analysis

Although they gave clear principles about collecting the views of the stakeholders, Pawson and Tilley (1997) were less clear on how to aggregate the data. In their example on post-code marking of goods (section 3.1.5), they stated that the researchers interviewed the stakeholders and then gave the conclusion that it was the police presence and the accompanying publicity rather than the postcode marking that led to crime reduction, but there was no detailed analysis about how the conclusion was reached. The researchers might have analysed the interview data extensively or they might have come to an insightful understanding, which led to their conclusion. I also looked at Gill and Turbin's (1999) realist study of the impact of closed circuit television (CCTV) in stores. They interviewed 480 customers and 38 shop thieves. They summarised the relevant qualitative data under each hypothesis and concluded that CCTV should be considered more of a tool to help combat shop theft rather than a solution. CCTV encouraged staff to approach suspected shop thieves and it helped staff to monitor suspicious individuals. Perhaps more significant was their comment that staff were not only part of the mechanism through which

CCTV achieved a result but were also considered part of the context in which it was expected to work. Thus, staff attitudes and involvement with the system became far more important as staff awareness of security issues increased. Furthermore, they commented that while it was relatively easy to propose plausible CMO configurations, it was much harder to collect useful data for each of the three categories (context, mechanism and outcome). This study illustrated some of the issues with analysing *CMOs* in realist studies, which I alluded to in (section 3.4) such as mechanisms being categorised as context and vice versa and the difficulty with connecting data under three categories.

Finally, I examined the analytic style used in realist medical education papers for further insight on data analysis. Ogrinc and Batalden's (2009) used realist evaluation in the evaluation of teaching improvement of care to internal medicine resident physicians. While they gave some plausible mechanisms and contexts as well as described the outcomes (learner satisfaction and improved patient care as represented by the increased percentage of patients who received influenza vaccine on discharge), there was lack of detail about the data collected and their analysis. Hollenberg and colleagues' (2009) realist evaluation of a large multi-institutional interprofessional education (IPE) initiative were more detailed in their analytic description as they used an inductive thematic analytic approach to code their 142 interviews and generated themes. They admitted to using an open-coding method without reference to a pre-existing framework or a specific theoretical perspective which was strange as realist evaluation is a theory-driven inquiry. It was therefore unsurprising that, despite identifying some contexts and outcomes, they failed to

explore mechanisms (as all six IPE projects were listed as mechanisms) and ultimately failed to generate a CMO theory.

The knowledge gained through the examples above highlighted the importance of avoiding the potential pitfalls of qualitative analysis such as weak or unconvincing analysis, mismatch between the data and the analytic claims made, mismatch between theory and analytic claims, or between the research question and the form of analysis used (Braun & Clarke, 2006). Therefore, for guidance on how to interpret qualitative data, I looked to other authors and selected Maxwell (2013) and Miles et al. (2014). Maxwell (2013) clearly said,

> ***"There is no cookbook or single correct way for doing qualitative analysis; your use of these strategies need to be planned (and modified when necessary) in such a way as to fit the data you have, to answer your research questions, and to address any potentially serious validity threats to your conclusions." (Maxwell, 2013, p. 105)***

Maxwell (2013) divided his analytic options into three main groups: (1) memos, (2) categorising strategies (such as coding and thematic analysis) which are based on similarity, relations or resemblance, and (3) connecting strategies (such as summarising) which are based on contiguity. Similarly, Miles et al. (2014) recommended, in essence, three parts to the process of collecting and analysing qualitative data: reducing the data, analysing the data, conclusion drawing and verification. Miles et al. (2014) further noted that, when all the data comes from within a bounded framework (in this study, the boundary is described by FD activities on teaching), it could be analysed together even though the data was collected from different sources.

I already used memos to capture, as well as to facilitate analytic thinking about my data including reflection on my goals, methods, theory, experiences and relationships with participants. Hence I did not consider memos a separate analytic group and I focused my qualitative data analysis on categorising and connecting strategies. Categorising was done by coding the appropriate data under each category of context, mechanism and outcome while connecting strategy seek to establish the relationships and influence of one on the other (*CMO*) and relate this to the hypotheses of the inquiry. Explaining connections requires researcher's insight, interpretation of the data, and exploration of the underlying ideas, assumptions and conceptualisations that are theorised as shaping the data (Braun & Clarke, 2006). The application of these categorising and connecting strategies is described in detail in Chapter 5.

## 4.4 Summary

In this chapter, I have discussed the underlying methodology and strategy for data collection from the stakeholders, the data collection strategy for each phase of the research and the data analysis strategy. This was a theory-driven inquiry (albeit a middle-range theory) hence the mixed method data collection and the data analysis strategies were theory-driven (i.e. to support, modify or refute the hypotheses of the inquiry). In the next chapter, I will give a clear description of the methods used for each phase of the study and details of the data analysis used.

# Chapter 5
# METHODS

## 5 Introduction

In the last chapter, I explained the rationale for the methods chosen. In this chapter, I will describe the methods used for the purpose of collecting rich data, which could support or refute the hypotheses. In the first half of the chapter, I describe the process of ethical approval and then go on to describe the method chosen for each phase of the study. Phase I was the FD webpage data from all UK medical schools, Phase II was the data from the observations and interviews of educators attending a FD course at Warwick Medical School (WMS) and Phase III was the data from the interviews of educators and FDC at eight other medical schools. In the second half of this chapter, I describe the data interpretation and the analytic methods chosen.

When discussing the design of qualitative research, Miles et al. (2014) stressed that the collection of data should not be random but purposive to enable researchers satisfy their specific needs in a research study. My need in this study was to collect rich data from different stakeholders (based on their expertise) that would support the appropriateness of the hypotheses. Rich data are data that are sufficiently detailed and varied to provide a full and revealing picture of what is going on, and of the processes involved (Maxwell, 2012). In addition, Becker (1970) claimed that rich data,

*"counter the twin dangers of respondent duplicity and observer bias by making it difficult for respondents to produce data that uniformly support a mistaken conclusion, just as they make it difficult for the observer to restrict his observations so that he sees only what supports his prejudices and expectations." (Becker, 1970, p. 53)*

## 5.1 Ethical approval

When obtaining permission for the use of information on individuals it was important to follow ethical guidelines, hence I applied for and obtained ethical approval from the Biomedical Research Ethics Committee of the University (Appendix 3a). I was aware of my responsibility to participants particularly in relation to the observations and interviews. To clarify this, I wrote participant information leaflets for the observation as well as for the interviews (Appendix 3b), which described the project, what participants had to do, and how the data would be used. If they were happy to participate, they were asked to sign the consent form. Once the data were collected and analysed, I destroyed all personal data to take care of confidentiality, anonymity and data protection issues.

## 5.2 Phase I data collection

I reviewed FD activities on the webpages of all 33 UK medical schools (Table 2.3) listed on the Medical Schools Council website (Medical Schools Council, 2010a). Initially, I carried out a scoping exercise looking at the webpage of every 5$^{th}$ medical school listed alphabetically. Information collected included medical school entry details, types of FD activities, details, duration / location of FD activities, fees, details of the FDC and HEA accreditation of FD courses. Furthermore, I noted the Complete University Guide (CUG) ranking of the medical schools. In UK, the CUG is used to

provide ranking of university courses (subjects) and generate a subject league table. The data for the subject league tables are the same as for the main university league tables, except that only four measures are used: student satisfaction, research assessment, entry standards, and graduate prospects. This would be discussed further in Chapter 6 (Phase I findings).

The HEA is the national body for learning and teaching in higher education and its accreditation scheme provides a national professional benchmarking that reflects best practices as well as external confirmation that the institutional provision is aligned with the UK Professional Standards Framework (UKPSF)[3]. The UKPSF (glossary, pg. viii) framework is based on four descriptors and the HEA uses those four descriptors to provide a recognition and accreditation service as a Fellow of the HEA at four levels (associate fellow, fellow, senior fellow and principal fellow) as detailed in Appendix 3c.

As shown in Tables 5.1a and 5.1b, there was detailed information on FD activities on the webpages of the six selected schools. I also did a comparison of the FD information available on WMS public and secure webpages to see if there was a major difference between the two pages of which there was none (Table 5.2).

[3] UKPSF: www.heacademy.ac.uk/ukpsf

**Table 5.1a: Scoping information on FD from six medical schools websites (2012)**

| Medical School | Entry details | Ranking (CUG) | FD Activities | Details | Duration | Location | HEA Accreditation | Course Fees | FD Co-ordinator |
|---|---|---|---|---|---|---|---|---|---|
| **Bristol** | Undergraduate<br>Graduate<br>Premedical entries<br><br>Total no of places 245 / year | 19th | Short Courses | 'Fit2Teach' Programme<br><br>'Professional and Generic Skills Programme' (PGSP) | ½ day workshops<br>1 day workshop | Various locations | | Free<br><br>Free | |
| | | | Teaching and learning for Health Professionals<br>*Teacher Education Funding from SHA* | PGCE (60 credits)<br><br>PG Diploma (120 credits)<br>MSc (since 2006) dissertation<br>Total 180 credits | 2 days +<br>I day induction | Centre for Medical Education<br>Medical school | Standard descriptor<br>1 (30 credits)<br>2 (60 credits) | £740 / 15 credits<br><br>£2940 | Dr Wendy Peek<br>Teaching Unit Developer<br>epzwjs@bristol.ac.uk<br>(0)117 33 11869/845 |
| **Edinburgh University** | Undergraduate<br><br>Total no of places / year not provided | 2nd | Teaching, Learning and Assessment (TLA)<br>*N.B: Contractual for new teachers; nested within the PGCert.* | PG Certificate in University teaching<br>60 credits | 2 day orientation +<br>2 core courses and<br>2 optional modules | Central | Standard descriptor 2 | Free | Carolin Kreber Programme Director, PG Cert.<br>0131 651 6668<br>carolin.kreber@ed.ac.uk<br>Daphne Loads - Head of Academic /Teaching Staff<br>daphne.loads@ed.ac.uk |
| **King's College School of Medicine**<br><br>5 entry routes to the MBBS program, the most of any UK medical school | Undergraduate<br>Graduate<br>Professional entries for Medicine.<br>For Maxillofacial<br>Direct entry<br>Extended medical degree<br>Total no of places / year not provided | 17th | Short courses<br>E-learning modules | | ½ day workshops | Guys campus | | | |
| | | | Professional Courses | Professional Certificate in Teaching and Learning in Higher/ Professional Education (60 credits<br>MA in clinical education<br><br>MA/PG Dip/PG Cert | 1-2 years<br><br>1yr Full time or 2 years PT | Central at the Institute of Education | Standard descriptor 2 | <br><br>£4338 | Course Administrator<br>Alison Finlay<br>a.finlay@ioe.ac.uk<br><br>020 7612 6362 |
| | | | Clinical Pedagogy | Graduate Certificate in Academic Practice (GCAP)<br>PGC / PGDipl /MA in Academic Practice | 1-2 years<br><br>1yr Full time or 3 years PT | King's Learning Institute | Standard descriptor 1<br>Standard descriptor 2 | £315/15 credits<br><br>£6650 | Learning and Teaching Coordinator: Dr<br>helen.graham@kcl.ac.uk<br>02078483905 |

**Table 5.1b: Scoping information on FD from six medical schools websites (2012)**

| Medical School | Entry details | Ranking (CUG) | FD Activities | Details | Duration | Location | HEA Accreditation | Course Fees | FD Co-ordinator |
|---|---|---|---|---|---|---|---|---|---|
| **Manchester Medical School (MMS)** | Undergraduate<br><br>Premed entry<br><br>Total no of places<br>400 / year | 12th | Walport Integrated Academic Training | Designed to facilitate the development of clinician academics of the future | | NHS partners | None | Free | Valerie.Wass@manchester.ac.uk |
| | | | Short courses/ workshops | New academics programme<br><br>1-3 years Part Time | ½ day workshop | Central | Standard descriptor 2 | | Faculty Training Manager<br>judith.c.williams@manchester.ac.uk<br>01612751468 |
| | | | Work Based Medical Education<br>*N.B: Available to internal candidates already teaching at the University and selected by the medical school only* | PGCE (60 credits)<br><br>PG Diploma (Total 120 credits)<br><br>MSc (dissertation) Total 180 credits | 5 two day sessions: 10 days<br><br>5 two day sessions: 10 days | Medical school | Standard descriptor 1 | £1,667<br><br>£ 1,667<br><br>£ 1,667 | Dr Don Bradley<br>0161 275 1851 |
| **Queen's University Belfast (QUB)** | Undergraduate<br><br>Total no of places<br>262 / year | 26th | Postgraduate Taught courses | PGCert, Dipl, MSc in clinical education (60/12/180 credits) | 1-3 years PT | Medical school | | £1500/ £1500/ £1200 | Mairead Boohan<br>m.boohan@qub.ac.uk<br>02890975068 |
| | | | PGCHET<br>*N.B: New lecturers on probation get priority* | Post-Graduate Cert. in Higher Education Teaching (60 credits) | 1-3 years | Centre for educational development | Standard descriptor 2 | Free for QUB staff | Linda Carey<br>l.carey@qub.ac.uk<br>028 9097 6610 |
| **Swansea** | Graduate<br><br>Total no of places<br>70 / year | Not ranked | THE Scheme<br>*N.B: Requirement for new teaching staff on probation from Sept 09*<br><br>Postgraduate courses | Teaching certificate scheme<br><br>PGCert., Diploma, Masters Teaching in Higher Education | Serial ½ - 1 day workshop<br><br>1 -3 yrs. FT<br>2- 5 yrs PT | Central at the Staff Development Unit<br><br>Central | Standard descriptor 2 | Free | Andrew Morgan, Head of Staff Development<br>a.j.morgan@swansea.ac.uk.<br>Ext. 5962. |

**Table 5.2: Scoping information on FD at Warwick Medical School website with and without secure access (2012)**

| Medical School | Entry details | Ranking (CUG) | FD Activities | Details | Duration | Location | HEA Accreditation | Course Fees | FD Coordinator |
|---|---|---|---|---|---|---|---|---|---|
| **Warwick Medical School (WMS)** | Graduate entry only | 14th | **FD INFORMATION WITHOUT ACCESS** | | | | | | |
| | | | Short courses | Postgraduate Award (PGA) Essentials of Clinical Education (20 CATS) | 3 days | Medical school | Standard descriptor 1 | £895 (Free to teachers) | Teaching and Learning Specialist Catherine.Bennett@warwick.ac.uk Tel: 024761 50869 |
| | Total no of places 178 / year<br><br>Largest number of graduate entries in UK | | Postgraduate Taught courses<br><br>*N.B: 50 – 100% bursary available to people teaching on behalf of WMS and Senior WMS employed teachers* | PGCE (60 CATS)<br><br>PG Diploma (Total 120 CATS)<br><br>MMed Ed (Total 180 CATS) | 2-8 years Part Time | Medical school | Standard descriptor 2 | £2,685<br><br>£2,685<br><br>£2,685 | Teaching and Learning Specialist Catherine.Bennett@warwick.ac.uk Tel: 024761 50869<br><br>Course Director Masters in Medical Education Neil.Johnson@warwick.ac.uk Tel: 02476 574573 |
| | | | **FD INFORMATION WITH ACCESS** | | | | | | |
| | | | Same as above | | | | | | |

Having established that there was information on FD on the webpages of the medical schools, I piloted the webpage scoring index on two postgraduate deaneries websites (London and West Midlands) to check for consistency and accuracy. The webpage scoring index as described in section 4.2.1.1 had three broad categories and 11 criteria (Table 4.1) with each criterion scored 0, 1 or 2 (maximum 22 points). The pilot data showed the 11 criteria to be consistent and accurately reflected the component of the webpage that it was applied to. The index was then applied to the webpages of the 33 medical schools, which were classified as follows: score 0-6 poor webpage, 7-12 average, 13-17 good and 18-22 excellent webpage.

I also collected available CMO data on FD from the medical school webpages to help refine and modify the hypotheses. I used thematic analysis to extract this data. Braun and Clarke (2006) have defined thematic analysis as,

> ***"Identifying, analysing and reporting patterns (themes) within data. It organises and describes the data set in (rich) detail. However, frequently it goes further than this, and interprets various aspects of the research topic." (Braun and Clarke, 2006, p.79)***

Braun and Clarke (2006) went further to explain that thematic analysis involves searching across a data set to find repeated patterns of meaning (themes – glossary, pg. viii). I used a theory-driven deductive thematic analysis approach to extract CMO data from the webpages. The categories were already decided by the CMO hypotheses as context, mechanisms and outcomes. Coding of webpage textual data was done using the low-inference (manifest) approach (section 4.2.1.1) to produce themes, which remain 'close' to the primary textual data. For example under the

category of context, themes included ‘school policy on FD’ and ‘institutional value of teaching’ with relevant data coded underneath each theme. The findings are discussed in Chapter 6.

## 5.3 Phase II data collection

This was a ‘non-participant observation with interaction’ of a FD activity at WMS, aimed at understanding the culture and process of FD in a medical school context (setting and objectives), observing FD (delivery and methods) and using interviews to explore the views of participants on FD (reception and perception). I used the postgraduate award ‘Essentials of Clinical Education’ (ECE) course, a three day programme designed for up to 25 participants, held three times a year. The course has 24 hours contact time (8 hours daily), and 175 hours of self-directed learning supported by an online discussion forum for each cohort. Details of the course and the post course assessment are shown in Table 5.3. I chose this course as it was accredited by the HEA at an associate fellow level as described in section 5.2. The ECE course is the first module for the postgraduate certificate, diploma and masters in medical education hence there would be a variety of educators attending the course who would still be contactable six months later. The aim of the course as stated on the WMS website[4] was,

> ***“To develop the understanding of education in the field of clinical practice and to begin to apply that understanding in the workplace.”***

[4] ECE course aim on WMS website: www2.warwick.ac.uk/fac/med/study/cpd/current/pgle/modules/md962

Participants' views were explored by immediate informal interviews during the course and six months later by formal semi-structured interviews. Informal interviews are casual conversations with informants (without the use of a structured interview guide) while in the field as an observer (Bryman, 2012). I used jottings and brief notes to help in the recall and writing of interview notes. Semi-structured interviews are formal interviews with the use of an interview guide (a list of questions or topics that need to be discussed during the conversation) but as the researcher, I was able to follow topical trajectories in the conversation that might stray from the guide when I felt this was appropriate (Bryman, 2012).

**Table 5.3: Details of the Essentials of Clinical Education Course (ECE)**

| Day 1 | Day 2 | Day 3 |
|---|---|---|
| **0900 – 1630** | | |
| Getting to know you: Welcome, Introduction & Overview | On the Job Teaching | Teaching and learning in large groups |
| Being a clinical educator | E-Learning | Portfolios for learning and assessment Includes guidance on module assessment |
| How to teach a practical skill | Planning a teaching session | Facilitating learning in small groups |
| Research and scholarship in clinical education | Strategies for active and interactive learning | What next? Developing and evidencing teaching practice Includes portfolio assessment Q&A session |
| Assessment, evaluation and feedback | Reflection and feedback on day 2 | Reflection and feedback on day 3 |
| Reflection and feedback on day 1 | | Module evaluation session |
| | | Close |
| **Post Course Assessment** Assessment is based on a reflective teaching portfolio comprising seven key tasks: | | |
| Teaching observations: Two teaching observations with reflective comments (one as teacher, one as observer) Giving feedback to learners: Reflective comments on giving feedback to students on at least three occasions. Evaluation: Evidence of collecting and using feedback from students. Lesson planning: Ensuring effective learning in clinical settings. Appropriate use of technology to enhance learning: Review of a learning event. Commitment to the HEA Professional Values Statements: A reflective overview. Action Plan: Future professional development as an educator. | | |

Phase II data collection was carried out in three steps as described below.

*Step 1*: *Interview of WMS FDC*

I carried out a one-hour in-depth, semi-structured interview of the WMS FD co-ordinator on the range of FD activities available, participation, outcomes and the FDC views on the benefits of FD at individual and institutional levels. This was followed by specific questions on the ECE course including the aim / objectives, intentions, evolution of the course, evaluation / outcomes and, more importantly from a realist perspective, what were the possible explanations (mechanisms) for the observed outcomes. This interview was taped and transcribed for analysis prior to observation of the two courses.

*Step 2: Observations and Interviews during the course*

Prior to the ECE course, I developed an observed engagement scale as discussed in section 4.2.2.1. This was a 5-point engagement scale where 1 represented no engagement and 5 represented full engagement (Table 4.2). I also designed an observation schedule (Appendix 3d) to record individual and group data as well as keep track of participants during the sessions. The observed engagement scale and the observation schedule were piloted on a different course prior to the ECE course. The course used for piloting (Becoming an Effective Teacher) runs for five days at WMS but I observed the participants over a three day period to simulate the ECE course. This allowed me to test my ability to listen, observe and take notes during the sessions and at break times. I was able to test my observed engagement scale and fine-tune my observation template prior to observing the ECE course.

During the actual ECE course, I observed, listened, wrote field notes and carried out interviews of participants during each course in January and April. The observations focused on individual participant as well as the interaction amongst participants (for example at their tables, during tasks and discussions), but the key focus of the observation was on participants. Observation data included session organisation, number of attendees, topics covered, teaching methods, evaluation, facilitators and specialist areas, discussions during sessions and breaks. In other words, data were collected on the physical setting as well as individual and group interactions. Each session was about one hour long and with 22 to 24 participants sitting in groups of threes and fours, observation of each table was carried out at five minutes intervals thereby providing two sets of observations for each individual per session.

I also scored participants' engagement on the observed engagement scale (Table 4.2); the scale was not given to participants. Engagement was scored using the 5-point scale. When all three categories descriptors (behavioural, emotional and cognitive) were observed a score of 5 (full engagement) was given, 4 (mostly engaged) if any two categories descriptors were observed, 3 (partial engagement) if only one category descriptor was observed, 2 (minimal engagement) if only one category descriptor was observed only once during the two sets of observations and a score of 1 (no engagement) if no category descriptor was observed. The scores for individual participants were recorded session by session over the three day period to document the trend in individual engagement over the three days, as well as to determine if there was a group effect (e.g. low engagement) for any particular session. In addition, participants were asked at intervals (breaks, lunch and end of

the day) to describe their engagement during the sessions (as discussed next under interviews).

Participants were interviewed during the course to explore reasons for their attendance and their opinions about the course relevance, usefulness, applicability and attributes of FD. These informal interviews were carried out in order to add further information to the observations and in order to establish a foundation for a deeper understanding of what had been observed (Hammersley & Atkinson, 2007). I chose informal interviews of participants at this initial stage as they are carried out 'on the fly' and do not require scheduling or taping. In fact, respondents might just see this as 'conversation', thereby fostering 'low pressure' interactions and allow respondents to speak more freely and openly (Cohen & Crabtree, 2006). It was helpful in building rapport and provided understanding from the participants' perspectives. Furthermore, according to Cohen et al. (2011), there would be candour, richness, authenticity and honesty about participants' experiences. Demographic data including gender, seniority, funding source, job type and previous attendance at a teaching course were also collected from all participants. Observation notes and notes from informal interviews were transcribed after each day. Observing and interviewing two groups allowed comparative analysis, improved validity and as observation is still uncommon in this area of FD research, it provided interesting and valuable data.

The key purpose of this step two of the study was to collect immediate data through observations / short informal interviews on the three common mechanisms

(motivation, engagement and perception) reported in the literature as being important in a learning process (section 2.9.2). These three mechanisms were used to develop some of the hypotheses listed in Table 3.2. I was aware that this would not explore all the possible hypotheses listed in the table; the others were explored with further data collection described in step three below.

*Step 3: Interviews of Course Participants at Six Months*

Six participants from each course were selected for follow up interviews six months later (total of twelve), which was within the recommended qualitative methodologic standard of at least eight participants for in-depth interviews (Kennedy et al, 2004; McCracken, 1988). Purposeful selection was based on the engagement score (above average, average and below average), seniority and gender. The audiotaped interviews were one hour, in-depth interviews to explore the longer-term impact on learning / behavioural changes i.e. what and how the interviewees have changed, and what facilitated or hindered that change (Appendix 3e). This allowed me to understand causal mechanisms and their outcomes.

The interviews followed the principles of 'theorising the interview' (Pawson, 1996) (section 4.2.3). I explained the hypotheses so the educators could understand them. We then discussed the hypotheses using information from the FD course as well as information from their experience of teaching during the six months following the course. As I anticipated, the educators were able to work out their own ideas and hypotheses about the FD. I helped them develop their views through a process of checking that I had understood correctly, sometimes challenging what they said and

thus we engaged together in the process of conceptual refinement. The educators were generally positive about FD but there were also criticisms.

Once transcribed, all individual phase II data (observations and interviews) were sent to the relevant participant. This respondent or member validation (Bryman, 2012) was to avoid misinterpreting the meaning of what participant's said and did as well as the perspectives they had on FD. It was also an important way for me to identify my own biases and misunderstandings of what I had observed or heard. In addition, once transcribed and analysed, the data helped in refining the interview schedule for phase III as well as gave me clarity on the acceptability, understanding and importance of my questions. Further details of analysis are described in section 5.5.

## 5.4 Phase III data collection

In Phase III, I felt that a slightly different approach might help to gather more comprehensive or possibly different views. I decided to interview one FDC and one educator on the same day from each of the eight chosen medical schools. This was both for convenience and to compare and challenge opinion as this might provide further insightful data regarding the hypotheses of the inquiry.

I carried out in-depth, face-to-face, semi-structured, individual, one-hour interviews of one FDC and one medical educator at eight medical schools. The medical schools were selected by purposive sampling method as described by Patton (1990). The selection of the institutions was from the FD webpage categorisation result (excellent, good, average and poor) in phase I (Chapter 6) combined with other

parameters such as geographical location, newness of school, presence of FD activities and availability of interviewees (FDC and educator). Two medical schools were chosen from each FD webpage category of excellent, good, average and poor.

The FDC interviewed were responsible for running the FD activities in the medical schools (even though they had different titles at different institutions). The FDC were identified from the webpages (as described in Phase 1) and also by direct contact with the medical school. Once identified, the FDC were requested to provide details of educators with responsibility for teaching medical students, who had attended a FD activity within the last two years. The second name listed alphabetically on the list was chosen for interview (or the third name if second person listed was not available). Attendance of a FD activity within the last two years was chosen in view of recall and demonstrable impact.

Each interview followed the principles of theorising the interview. In the teacher-learner process we discussed the hypotheses and my realist ideas in order to help the FDC or educator understand the research project. In addition, I asked open questions about what had gone well, what had not gone well; impact (if any) on their teaching and how they knew things have changed. I asked further questions about education degree, participation in FD, types of FD activity available / attended, rationale for attending to achieve conceptual refinement. A copy of the interview schedule is included in Appendices 4a and 4b. Just as I did in phase II, once transcribed I sent the interview transcript to individual interviewee's for respondent validation. In addition, I compared my coding of the first three interview data with that

of an independent qualitative data analyst to check agreement on codes and themes. I found the themes were quite similar with similar data coded under the CMO.

## 5.5 Data interpretation

As explained in section 4.3, data analysis was theory-driven and involved a mixed method analysis of qualitative and quantitative data described in detail below.

### 5.5.1 Quantitative data analysis

#### 5.5.1.1 Descriptive

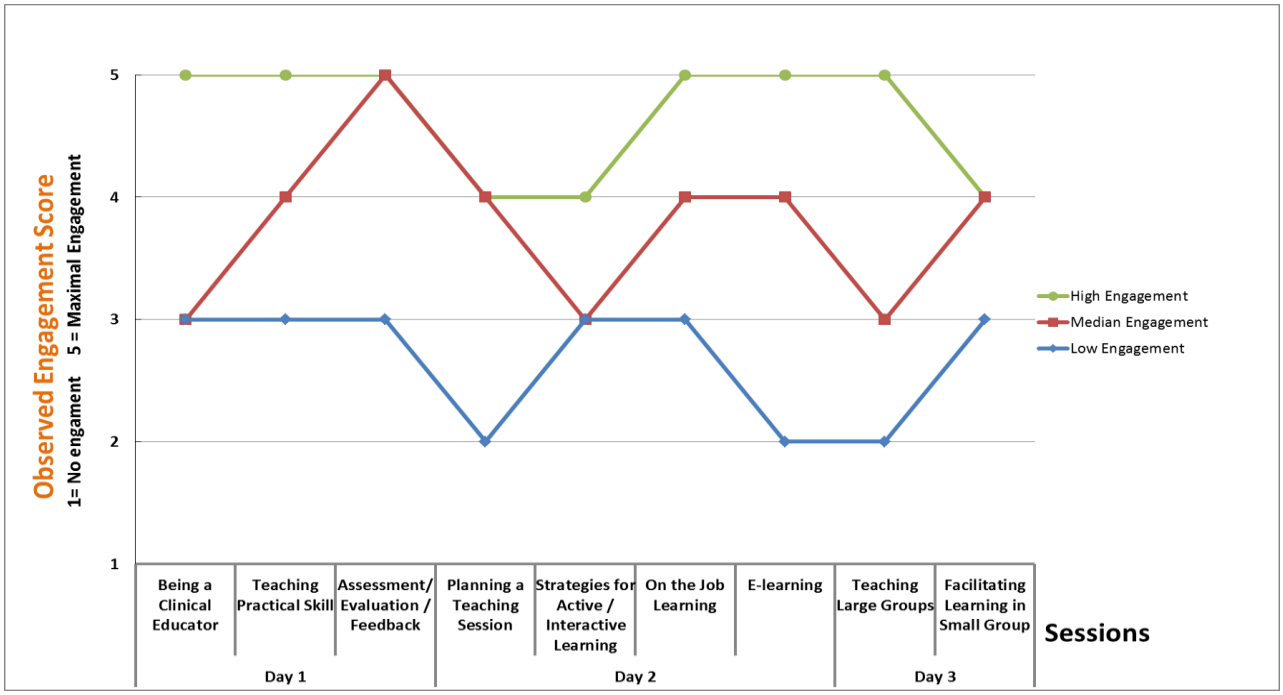
In phase I, the webpage scores were used to rank medical schools into four webpage categories (excellent, good, average, poor). Further comparative analysis was carried out based on medical schools categories such as student catchment, establishment date, geographical location and CUG ranking. This is discussed further in Chapter 6. In phase II, demographic data including gender, funding, previous teaching course attendance, seniority and job type were compared between the two cohorts (Jan and Apr) of medical educators that attended the ECE at WMS. The mean, median and range were calculated for the observed engagement scales of the two groups of educators.

#### 5.5.1.2 Slope diagrams

In phase II, the observed engagements (1 = no engagement, 5 = full engagement) were plotted sessionally on slope diagrams over the three day period for both

cohorts of educators. Slope diagrams were useful in giving a visual identification of outliers in terms of both absolute peaks and dips in engagement and it enabled me to compare trajectories to see if less engaged people have a different pattern from the more engaged (example in Fig 5.1 below). Course medians of the two cohorts were also compared and the engagement scores of the lowest engaged participant and the highest engaged participant were plotted against the median engagement of the groups for comparison.

**Figure 5.1: Slope diagram example**




### 5.5.1.3 Constructs

This was truly a mixed method approach as both quantitative and qualitative approaches were used to derive the bi-axial constructs. In phase II, the observations, engagement scores, and informal interviews of the 33 participants were used in developing bi-axial constructs for each of the three parameters (motivation,

engagement and perception) hypothesised to be important mechanisms of effective FD. The descriptors of the bi-axial constructs were derived by thematic analysis as described in section 5.2. Furthermore to ensure validity, I compared my coding of four randomly selected interview data with those of my three supervisors for coding agreement and to check that the descriptors of the constructs were derived from and reflected participants' data. Fisher's exact test was used to analyse the descriptors in each construct for statistical significance. The bi-axial constructs are described further in Chapter 7.

### 5.5.2 Qualitative data analysis

The first step in data analysis, the simplifying and abstracting of information, began during the observations and the interviews. In phase II, I wrote field notes for the observations and informal interviews but recorded the six months follow-up interviews. I recorded all the interviews in phase III and also wrote a summary of the data at the interviews. According to Emerson et al. (2011), the initial step in qualitative analysis is reading the observation notes, interview transcripts and documents to be analysed. Maxwell (2013) added to this and said, listening to the interview tapes prior to transcription and the actual process of transcribing or of rewriting and reorganising the rough observation notes are all opportunities for analysis. However, I was also aware of the dictum from Cohen et al. (2011) that any transcribing is inevitably interpretive and that transcribing has the potential for data loss and distortion. Therefore, soon after the observation or interview (usually on the same day), I read through and checked my observation or interview notes for

accuracy and then included any reflections I had about the observation or interview (for example, if the educator or the FDC was reluctant to discuss some aspects). By so doing, I hoped both to reflect the discussion accurately and also to have retained some aspects of the observation or interview as a social encounter.

I also considered whether the interviewees might not have been open in what they said and whether their responses might have been measured or guarded (e.g. by offering a corporate view instead of their own personal view). However, I found that both the educators and the FDC were confident and talked frankly and openly about the subject matter with authority. Therefore, I had no concerns that the data might be compromised by such factors as timidity, lack of honesty or lack of understanding on the part of the respondents. In addition, when I transcribed the interviews, I felt confident that generally their contributions were based on good levels of understanding of the topics they discussed. Once transcribed, all data were entered into the NVivo ($10^{th}$ edition) software package to aid analysis. Data analysis focused on using categorising and connecting strategies described next.

#### 5.5.2.1 Categorising strategy

This strategy is based on similarity. Similarity relations involve resemblance or common features; their identification is based on comparison, which can be independent of time and place. Similarities and differences are used to define categories and to group and compare data (Maxwell, 2013). Coding is a typical categorising strategy in qualitative research. The aim of coding is to order the data in an efficient data-labelling system so that it can be easily retrieved and worked upon

(Miles et al, 2014). The key codes and/or categories could be determined prior to the collection of data (Robson, 2011). The categories for this research project were the *CMO* – contexts, mechanisms and outcomes; they gave the basic structure to the data collection.

Once the main categories had been decided, I applied an eclectic coding style to the rest of the data. This involved the application of a combination of descriptive, evaluating and causation coding to segments of data, revising the code list as patterns in the data became more apparent as well as having 'conversations' with the data (Bazeley, 2007; Saldaña, 2013). This involved moving back and forth between the data sources, the codes, and the analytic framework (realist evaluation). With my realist lens, I iteratively formed assertions to capture the insights gained from the data about the impact of the programme on individuals, and more importantly what aspect works and in what context to refute or support the hypotheses.

During the observations and interviews, I wrote notes, comments and memos, which I later placed under headings of contexts, mechanisms and outcomes. As discussed in Chapter 3, there was some fluidity between contexts, mechanisms and outcomes and how data were categorised depended upon the realist hypothesis in which it was framed. These notes together with participants' data were arranged, in tabular form, under the appropriate hypotheses and sometimes the data provided information about more than one hypothesis. In addition, I was aware that some of the participants' data and comments might suggest completely new realist hypotheses.

#### 5.5.2.2 Connecting strategy

The sorting and coding of the data had used only the categorising skills of the researcher and so far I had made no complex judgements about the relevance of the material or made significant connections between the data hence the need for a connecting strategy. There are some descriptions of this approach in the literature but I found Maxwell's (2013) description most helpful,

> ***"Instead of fracturing the data into discrete segments and resorting it into categories, connecting analysis attempts to understand the data in context, using various methods to identify the relationships among the different elements of the text." (Maxwell, 2013, p. 112)***

Connecting strategy does not focus primarily on *similarities* but instead looks for relationships that *connect* statements and events within a context into a coherent whole (i.e. the influence of one on the other). A typical example of connecting strategy as suggested by Miles et al. (2014) is summarising the data in a visual and systematic display. When data are presented in this orderly way it can then be used for drawing conclusions. The difficulties in this process were ensuring that the summaries accurately reflected the data collected and also answered the purposes of the research (i.e. validating or disproving the hypotheses). Bearing in mind that the summary of the data had to depend on the researcher's judgement, I took very seriously Robson's (2011) advice on the deficiencies of the human analyst: I was careful not to ignore information that conflicted with my ideas, I took into consideration the reliability of the sources and I did not discount any information. I decided that the data should be presented as reflective summaries each headed by a hypothesis of the inquiry. Once the summaries were completed, in order to check

on the reliability of my judgement, I referred again to the raw data to ensure that my summaries in the data display reflected the views of the stakeholders.

## 5.6 Summary

This chapter culminates the first half of the thesis. In the previous chapters, I have reviewed the literature, decided on the realist framework, explored its methodology, and discussed in this chapter the methods for each phase of the project. The findings from each of the phases are presented in the next three chapters (Chapters 6, 7 and 8) with a relevant discussion in each chapter of the hypothesis supported or modified by the findings. However, a full discussion and the drawing of conclusions on the final theories that emerged will not be considered until Chapter 9 when the accuracy of the hypotheses will be assessed using all the data from the research.

# Chapter 6

# PHASE I FINDINGS

## 6 Introduction

In the previous chapter, I explained how information was collected about the webpages of UK medical schools related to FD. The rationale for this has already been discussed in section 4.2.1, which followed from the literature review that showed a paucity of publications on FD in UK medical schools. The question then arose as to how widespread FD provision is in UK medical schools and, if available, how the CMO data may be used to refine the realist hypotheses. This chapter focuses on discussing the findings from the webpage data analysis. This will be followed by a discussion of the CMO findings from a realist perspective and how the data support, modify or invalidate any of the hypotheses. A more detailed discussion of the findings from all phases is left till Chapter 9.

### 6.1 Webpages FD data

As discussed in section 2.1, various terminologies are used for FD including staff development, teaching / learning development, academic practice / organisational development and these were all used on the various webpages I reviewed. Similarly, the FDC had a variety of titles at the medical schools as explained in section 4.1.1. However, for simplicity, I have continued to use the terms FD and FDC. I reviewed FD webpages of 30 of the 33 UK medical schools (Table 6.1). I was unable to access the FD webpages of three medical schools (University of Cambridge School of Clinical Medicine, St George's Medical School and University of Glasgow Faculty

of Medicine) as their FD webpages were password protected in the staff intranet section. This raises some issues about these institutions perceptions of FD as to why FD activities are not visible, a point which I address in the discussion section.

Based on the scoring index criteria (section 4.2.1.1), seven (23.3%) medical school webpages were classified as *excellent* (scored 18-22), eight (26.7%) as *good* (scored 13-17), 12 (40%) as *average* (scored 7-12) and three (10%) as *poor* (scored 0-6). This meant half of the reviewed webpages were *poor* to *average* and the other half were *good* to *excellent*. There was no UK national, regional or geographic trend in the result of the webpage classification (Table 6.1).

Under general qualities category (Table 6.1), most webpages had problems with *currency* i.e. the extent to which material can be identified as being up to date, as half (15) of the websites had a score of 0 (they were dated more than twelve months) while more than one third (11) scored one as they were dated within six to twelve months but with no review date. Only four (13%) met the full criteria of *currency*. Under web utilisation category, the challenging criterion was *navigation* as one third (10) of the webpages scored zero because of issues such as poor layout, information organisation, lack of descriptive subheadings and difficulty moving back and forth between pages. Only eight websites (27%) fully met this criterion.

## Table 6.1: Medical Schools FD Webpage Scoring and Classification (NA – Not accessed)

| NAME OF MEDICAL SCHOOL | AUTHORITY | ACCURACY | CURRENCY | COVERAGE | OBJECTIVITY | FUNCTIONALITY | FLEXIBILITY | NAVIGATION | COMPREHENSIVENESS | LOCATION | CONTACT DETAILS | TOTAL SCORE | GRADE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| England | | | | | | | | | | | | | |
| **East Anglia** | | | | | | | | | | | | | |
| Cambridge (University of), School of Clinical Medicine | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | | |
| University of East Anglia | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 1 | **12** | **Average** |
| **Greater London** | | | | | | | | | | | | | |
| Barts and The London School of Medical and Dentistry | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | **16** | **Good** |
| King's College London School of Medicine (at Guy's King's | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | **19** | **Excellent** |
| Imperial College School of Medicine, London | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | **19** | **Excellent** |
| London School of Hygiene and Tropical Medicine (Postgra | 2 | 2 | 0 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | **15** | **Good** |
| St George's, University of London | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | | |
| University College London, University College Medical Sc | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | **20** | **Excellent** |
| **Midlands** | | | | | | | | | | | | | |
| Birmingham (University of), School of Medicine | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | **5** | **Poor** |
| Keele University, School of Medicine | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | **20** | **Excellent** |
| Leicester (University of), Leicester Medical School | 1 | 2 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | **12** | **Average** |
| Warwick (University of), Warwick Medical School | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | **12** | **Average** |
| Nottingham (The University of), Faculty of Medicine and | 0 | 2 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | **10** | **Average** |
| **North East** | | | | | | | | | | | | | |
| Durham (University of), Queens Campus, Stockton, Phas | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 2 | **11** | **Average** |
| Hull York Medical School | 1 | 2 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | **10** | **Average** |
| Leeds (University of), School of Medicine | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 0 | **15** | **Good** |
| Newcastle (University of), Newcastle Biomedicine, The M | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | **18** | **Excellent** |
| Sheffield (The University of), School of Medicine | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | **12** | **Average** |
| **North West** | | | | | | | | | | | | | |
| Liverpool (University of), Faculty of Health and Life Scien | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | **5** | **Poor** |
| Manchester (University of), Faculty of Medicine and Hum | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | **15** | **Good** |
| **South** | | | | | | | | | | | | | |
| Brighton and Sussex Medical School | 1 | 1 | 0 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | **14** | **Good** |
| Oxford (University of) | 1 | 2 | 0 | 2 | 2 | 1 | 0 | 0 | 2 | 2 | 0 | **12** | **Average** |
| Southampton (University of), School of Medicine | 2 | 2 | 0 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | **17** | **Good** |
| **South West** | | | | | | | | | | | | | |
| Bristol (University of), Faculty of Medicine | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | **20** | **Excellent** |
| Peninsula Medical School | 0 | 2 | 1 | 0 | 0 | 2 | 1 | 1 | 0 | 1 | 1 | **9** | **Average** |
| | | | | | | | | | | | | | |
| Scotland | | | | | | | | | | | | | |
| Aberdeen (University of), School of Medicine | 1 | 2 | 0 | 2 | 1 | 1 | 1 | 0 | 2 | 1 | 1 | **12** | **Average** |
| Dundee (University of), Faculty of Medicine, Dentistry ar | 2 | 2 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | **20** | **Excellent** |
| Edinburgh (The University of), The Faculty of Medicine | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | **16** | **Good** |
| Glasgow (University of), Faculty of Medicine | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | | |
| St Andrews (The University of), Faculty of Medical Scienc | 1 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | **6** | **Poor** |
| | | | | | | | | | | | | | |
| Wales | | | | | | | | | | | | | |
| Cardiff University, School of Medicine | 2 | 2 | 0 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 0 | **16** | **Good** |
| Swansea University, School of Medicine | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | **9** | **Average** |
| | | | | | | | | | | | | | |
| Northern Ireland | | | | | | | | | | | | | |
| Queen's University Belfast, Faculty of Medicine and Heal | 1 | 2 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 2 | **11** | **Average** |

In the FD specific category, according to the 30 webpages, most FD activities have been developed as a response to the needs of staff new to teaching, lecturers within their probationary period, experienced staff wishing to further develop specific areas of teaching, and other staff with significant roles in supporting student learning. However, the majority of the FD activities listed were formal programmes (e.g. postgraduate certificate with optional progression to diploma and master's degrees in education) or workshops. Presumably these were easier to describe and / or advertise; nevertheless it was similar to my literature review findings of courses and workshops being the most common FD types (Steinert et al, 2006). On the webpages, most of these courses did have costs attached with varying levels of reimbursement to faculty members.

Only one third of the webpages fully met the criterion on *comprehensiveness* with the other two-thirds distinctly lacking in their FD content description covering areas such as online learning and mentoring. Again whether this represents lack of provision, error of omission or provision only on password protected pages was difficult to ascertain. The location of the FD activity was another evaluation criterion based on the thinking that if the FD is located in-house in the medical school, it might be more tailored or focused to the need of the faculty teaching medical students. However, the distribution was almost equally spread as sixteen (53%) of the medical schools had their FD located centrally i.e. within the central university, meaning there was more FD provision related to higher education in general rather than medical

education specifically, while fourteen (47%) had at least some of the FD activities based within the medical school.

Details regarding the FDC or their equivalent were more challenging. Six medical schools had no FDC listed, which could be interpreted to mean they had none but this was not necessarily true as I subsequently found out. For example University of Birmingham School of Medicine and University of Liverpool Faculty of Health and Life Sciences had no FDC listed on their webpages, but I was later able to make telephone contact through the medical school and interview their FDC in phase III (Chapter 8). Of the other 24 schools, half (50%) had only a name and the other half had all the details including name, e-mail and telephone contacts. Understandably the contact details criterion is one which could be difficult to maintain accurately (as personnel or contact numbers change and departments merge) but ideally this is information that should be kept up to date and it raises questions about webmasters and dedicated web services which I address in the discussion section.

## 6.2 CUG ranking and FD webpages

Next, I looked at the CUG[5] ranking because if student satisfaction is used as a measure of the quality of teaching, then one would expect schools that are highly ranked (and value teaching) to have excellent webpages on FD opportunities. As mentioned in Chapter 5, the CUG is used to provide ranking of universities as well as a subject (courses) league table. It is one of three national rankings of universities /

[5] CUG: www.thecompleteuniversityguide.co.uk/league-tables/rankings

subjects published annually; the other two are published by The Guardian and The Times / Sunday Times. The CUG is published independently by the Mayfield University Consultants. The primary aim of ranking is to inform potential applicants about universities / subjects based on a range of quality criteria. The data used for the subject league tables are the same as for the main university league tables, except that only four measures are used: student satisfaction, research assessment, entry standards, and graduate prospects. All four measures are given equal weighting and an overall score is calculated. To qualify for inclusion in a subject table, a university has to have data for at least two of the four measures.

The raw data for the CUG league table come from sources in the public domain. The Higher Education Statistics Agency (HESA) in the UK is the official agency for the collection, analysis and dissemination of quantitative information about the universities. HESA provides data for entry standards, student-staff ratios, spending on academic services, facilities spending, graduate prospects, completion and overseas student enrolments. However, the data for student satisfaction, which is a measure of students' view on teaching quality, is provided by the National Student Survey (NSS)[6]. The NSS, launched in 2005, is a questionnaire survey of all final year degree students at UK institutions as part of a quality assurance framework. The aim is to gather feedback on course quality to make institutions more accountable for the experience they are delivering to students and to help inform the choices of future applicants to higher education. All universities are obliged to

[6] NSS: www.thestudentsurvey.com

provide contact details for eligible students even though participation is voluntary for the students themselves. The survey has a core set of 22 attitude questions designed to assess students' opinions of the quality of their degree programmes in seven areas of learning experience which are:

- Teaching on the course
- Assessment and feedback
- Academic support
- Organisation and management
- Learning resources
- Personal development
- Overall satisfaction

These are supplemented by open ended questions to capture any particular positive or negative aspects that the student wishes to highlight.

Interestingly there was no correlation between the CUG (2012) ranking of medical schools and the FD webpage scores (Table 6.2). In fact the distribution was quite random (Figure 6.1). There are possible explanations for this. One is that student satisfaction is only one of four criteria (weighted equally) to produce the CUG ranking; hence the other three criteria might act as confounders. Second is the suggestion that some universities advise students to artificially inflate the scores they give in the NSS in the interest of improving the university's ranking and by extension students' job prospects. One highly publicised example of this was at Kingston University where the lecturers allegedly told the students to give Kingston good scores because "if Kingston comes bottom ... no one is going to want to employ you because they'll think your degree is **** " (Swain, 2009). Apparently this was part of

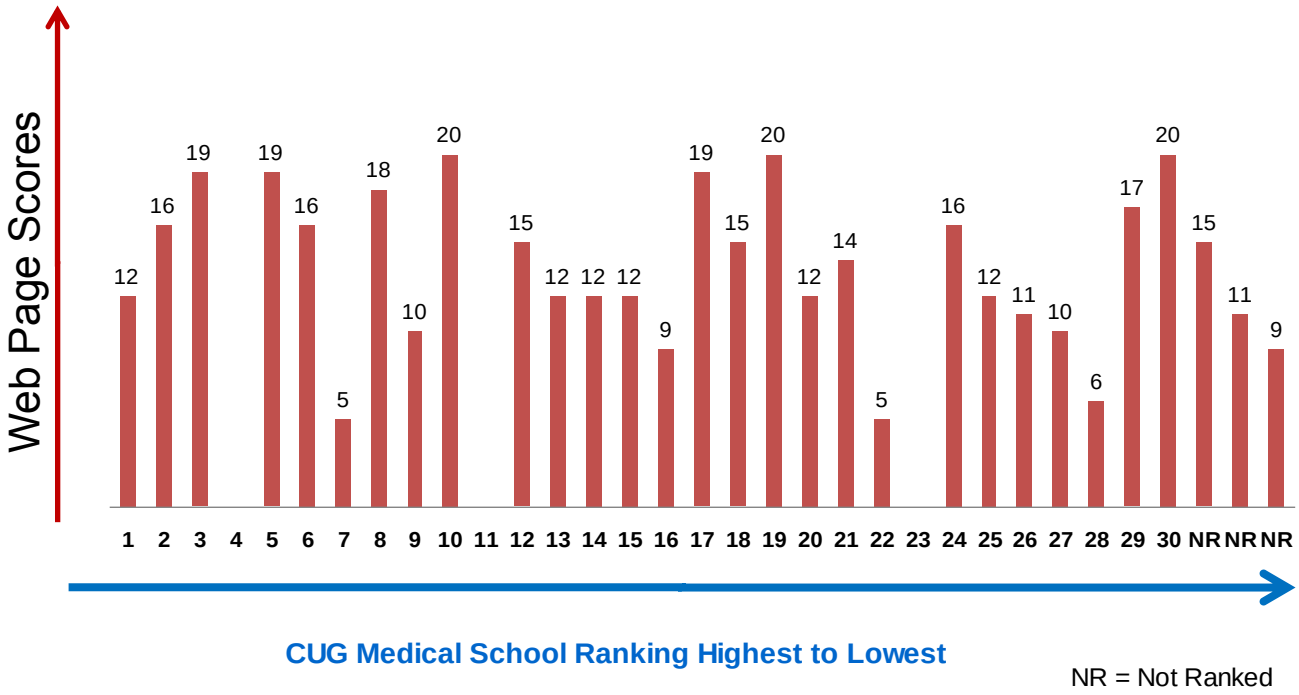
'reputation management' as lecturers warn students that prospective employers and postgraduate courses use the NSS to assess the value of their degree. Following an investigation of the allegations, the Higher Education Funding Council of England (HEFCE)[7] ordered that Kingston University's Department of Psychology be removed from the 2008/09 league tables (Coughlan, 2008).

**Table 6.2: Webpage scores and CUG ranking (NA – Not accessed)**

| NAME OF MEDICAL SCHOOL | CUG RANKING 2012 | TOTAL SCORE | GRADE |
|---|---|---|---|
| Oxford (University of) | 1 | 12 | Average |
| Edinburgh (The University of), The Faculty of Medicine | 2 | 16 | Good |
| University College London, University College Medical School | 3 | 20 | Excellent |
| Cambridge (University of), School of Clinical Medicine | 4 | NA | NA |
| Imperial College School of Medicine, London | 5 | 19 | Excellent |
| Barts and The London School of Medicine and Dentistry | 6 | 16 | Good |
| Birmingham (University of), School of Medicine | 7 | 5 | Poor |
| Newcastle (University of), Newcastle Biomedicine, The Medical School | 8 | 18 | Excellent |
| Hull York Medical School | 9 | 10 | Average |
| Dundee (University of), Faculty of Medicine, Dentistry and Nursing | 10 | 20 | Excellent |
| Glasgow (University of), Faculty of Medicine | 11 | NA | NA |
| Manchester (University of), Faculty of Medicine and Human Sciences | 12 | 15 | Good |
| Sheffield (The University of), School of Medicine | 13 | 12 | Average |
| Warwick (University of), Warwick Medical School | 14 | 12 | Average |
| Leicester (University of), Leicester Medical School | 15 | 12 | Average |
| Peninsula Medical School | 16 | 9 | Average |
| King's College London School of Medicine (at Guy's King's College and St Thomas' Hosp) | 17 | 19 | Excellent |
| Leeds (University of), School of Medicine | 18 | 15 | Good |
| Bristol (University of), Faculty of Medicine | 19 | 20 | Excellent |
| Aberdeen (University of), School of Medicine | 20 | 12 | Average |
| Brighton and Sussex Medical School | 21 | 14 | Good |
| Liverpool (University of), Faculty of Health and Life Sciences | 22 | 5 | Poor |
| St George's, University of London | 23 | NA | NA |
| Cardiff University, School of Medicine | 24 | 16 | Good |
| University of East Anglia | 25 | 12 | Average |
| Queen's University Belfast, Faculty of Medicine and Health Sciences | 26 | 11 | Average |
| Nottingham (The University of), Faculty of Medicine and Health Sciences | 27 | 10 | Average |
| St Andrews (The University of), Faculty of Medical Sciences | 28 | 6 | Poor |
| Southampton (University of), School of Medicine | 29 | 17 | Good |
| Keele University, School of Medicine | 30 | 20 | Excellent |
| London School of Hygiene and Tropical Medicine (Postgraduate Medical School) | Not ranked | 15 | Good |
| Durham (University of), Queens Campus, Stockton, Phase 1 Medicine | Not ranked | 11 | Average |
| Swansea University, School of Medicine | Not ranked | 9 | Average |

[7] HEFCE: www.hefce.ac.uk

**Figure 6.1: Webpage Score & CUG Ranking**




NR = Not Ranked

## 6.3 CMO data

In Chapter one I explained that contexts are the geographical and social framework of the programme under study; mechanisms come from the structure of the programme and are what makes the programme effective, while outcomes result from the programme. I was able to identify some potential contexts from the FD webpages as well as some intended outcomes for the FD activities (Table 6.3). The data were extracted by thematic analysis as described in section 5.2.

**Table 6.3: Contexts and Outcomes from FD webpages**

| Contexts | | Outcomes |
|---|---|---|
| *Infra-structural* | The need for medical teachers to be trained / become qualified in teaching<br>*Examples: Teaching certificate scheme is a requirement for new teaching staff on probation at* ***Swansea Medical School*** *from Sept 2009 onwards*<br><br>*Teaching, Learning and Assessment (TLA) contractual for new teachers; nested within the PGCE at* ***Edinburgh University Medical School*** | Improvement in instructional skills<br>*Increased knowledge of educational concepts and principles*<br>*Gains in skills (e.g. assessing learners' needs)*<br>*Promote reflection and provision of feedback to learners.*<br><br>Quote from a participant:<br>*"I think that my teaching practice has improved immensely over the time I have been studying on the TLHP course. I use the skills I have gained on a daily basis and I am keen to continue this development"*<br>***Faculty of Medicine, University of Bristol***<br>*www.bris.ac.uk/medical-education*<br><br>Obtain formal and recognisable qualification in learning and teaching.<br>***University of Aberdeen*** *certificate in medical education*<br><br>Sense of commitment to and increased enthusiasm to teaching<br>***University of Bristol****, Teaching and Learning for Health Professionals (TLHP)*<br><br>Career progression in teaching and further development as a teacher |
| *Institutional* | School policy on FD<br>*Example:* ***Manchester Medical School:*** *Work Based Medical Education is available to internal candidates already teaching at the university and selected by the medical school only*<br><br>Cost / resources for FD activities<br>*Examples:* ***Warwick Medical School:*** *Bursary is available to medical school teachers for postgraduate taught courses (PGCE)*<br><br>***University of Aberdeen****: There is no direct cost to teaching staff in medicine for the PGCert in medical education*<br><br>Institutional value of teaching<br><br>FD located in medical school or centrally | |
| *Individual* | Details of facilitators – title, level of expertise<br><br>Prior learning/reading/preparation before the FD<br><br>Time – number of hours to attend FD and post course assessment (including timetable)<br><br>Positive experience of FD from previous participants quotes on webpages<br><br>*"The 5-day course proved stimulating and challenging. I will advise my colleagues to attend"*<br>Participant quote post attendance at the ***University of Dundee discovery courses in medical education*** | |
| *Inter-personal* | Participatory, multimodal approach to teaching<br><br>Blended learning approach<br>*Example: 4-6 contact days/module and facilitated e-learning to maximise flexibility at* ***Peninsula Medical School***<br><br>Shared needs / addressing relevant areas of educators needs. Learning with peers | |

Most of the potential contexts and possible outcomes were quite similar to the ones already identified in my literature review (section 2.9). While it is difficult to be certain if these intended outcomes were achieved, some schools provided participants quotes related to outcomes on the webpages. I was not able to infer mechanisms from the webpage analysis, which was perhaps unsurprising, as I have chosen a *low-inference* (manifest) analysis (Maxwell, 2013) to achieve better reliability as discussed previously in section 4.2.1.1. Nevertheless, the context and outcome data obtained supported some of the hypotheses of the inquiry listed in Table 3.2. The hypotheses mostly supported were hypotheses 2, 3 and 6 as shown in Table 6.4. More importantly from a UK perspective, the findings were similar to those reported in the literature review; hence there was no need to change the hypotheses at this stage of the research.

**Table 6.4: Hypotheses supported from webpage data**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 2 | FD course aligned to the needs of the educator or FD relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills and change in teaching approach |
| Hypothesis 3 | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Increased enthusiasm, interest and higher importance attached to teaching |
| Hypothesis 6 | Outside initiatives / external influences and demand to standardise medical educators teaching | + | Professionalization of teaching with qualification / accreditation and standards by recognised bodies or authorities | = | Career progression in teaching / improved credibility and recognition as educators |

## 6.4 Discussion

There were three key outputs from the webpage analysis, which were as follows:

i. It provided evidence of the provision of FD activities in the 30 UK medical school webpages that could be accessed
ii. It contributed to the realist framework by providing some contextual data and intended outcomes
iii. It provided a classification template (*excellent, good, average, poor*) to be used in phase III selection of medical schools.

The first key output of the webpage analysis was on FD provision. The webpages answered the question of how widespread FD provision is in UK medical schools. This question arose from the literature review finding of paucity of UK publications on FD as reported in Chapter 2. Hence, it was interesting to find FD opportunities described on all the 30 UK medical school webpages accessed. However, it did not inform the second part of the question as to why the small number of publications. There are many possible reasons for lack of engagement in educational scholarship including competing demands, time pressures, lack of motivation, lack of rewards for publication, lack of skill / lack of support in educational research and the importance attached to educational scholarship (Goldszmidt et al, 2008; Zibrowski et al, 2008). These reasons were explored further with the FDC in phase III to identify those pertinent to UK and the findings are reported in Chapter 8. Here though, I discuss the importance of FD education scholarship in general.

The issue of education scholarship was addressed by Boyer (1990) when he categorised the four scholarship areas: discovery, integration, application, and teaching. The first category, the scholarship of discovery is often synonymous with research in the traditional sense and potentially results in peer-reviewed publications and grants. The second category, the scholarship of integration, has been defined as making connections across the disciplines and illuminating data in a revealing way, whereas the third category, the scholarship of application, has been likened to service in one's own field of knowledge and the application of theory into practice. The scholarship of teaching involves the capacity to effectively communicate one's own knowledge and skills. It has also been said that teaching becomes scholarship when it is made public, is available for peer review, critique and can be reproduced and built on by other scholars (Beattie, 2000; Rice, 2002).

FD in medical education has to engage in scholarship for it to be refined and built on by other scholars, improve the education process, improve education research productivity, demonstrate impact and cost effectiveness (Goldszmidt et al, 2008; Gruppen et al, 2003). There have been repeated calls for scholarship in FD and for active support from the institutions (Whitcomb, 2002; 2003; Gruppen et al, 2006). Currently there is no doubt that there is growing pressure for medical teachers to have a scholarly approach to teaching as evidenced by the latest GMC document document on the 'recognition and approval of trainers in the UK' (GMC, 2012), and the Association of American Medical Colleges (AAMC) document on the 'evidence of educational scholarship' (AAMC, 2007). It is good that medical schools are responding by making more FD opportunities available but the issue of FD scholarly

activity in UK medical schools still needs to be addressed and hopefully my study will help towards addressing this issue.

The second output from the webpage review was the extraction of *CMO* data to reflect a UK perspective on FD and refine the realist hypotheses. I was able to identify some contexts and outcomes from the webpages. Contextual data were perhaps the easiest to extract from the webpages as most schools provided some details of their FD opportunities. Outcomes were a bit more challenging to extract as quite a lot were written as objectives i.e. in terms of OBE (section 2.4.1) rather than as realist outcomes. Furthermore, the outcomes have to be viewed as potential or intended outcomes as only a few webpages (three) provided quotes from participants to corroborate achievement of these outcomes.

Mechanisms are the agents of change: they describe how the structures and resources embedded in a programme influence the reasoning and behaviour of the programme subjects. Hence mechanisms require an evaluative approach of programme participants, which would be difficult to infer from webpages. Furthermore, as explained in section 4.2.1.1, I chose a *low-inference* (manifest) approach for the webpage analysis, as this was important for a realist approach and to achieve a more reliable result. Therefore, it was perhaps unsurprising that I was unable to infer any mechanism from the FD descriptions on the webpages. Moreover, I bore in mind the limitation of textual analysis inability to portray a rich understanding of meaning and causation (Graneheim & Lundman, 2004). Nevertheless, there was still enough data to support three of the hypotheses.

The third output of the webpage analysis was the classification of the medical schools into four groups based on the eleven criteria discussed in Chapter 5. These became a purposeful sampling template for phase III. However, as discussed under the findings (section 6.1), there was no correlation between the CUG rating and the FD webpage classification even though the expectation was that the highly rated medical schools would have excellent FD webpages for their teachers and learners. Possible confounders made this interpretation less straightforward. Furthermore following the interviews at the eight medical schools, I will explain in Chapter 8 how similar the findings on FD were from all the schools despite the classification.

Nevertheless, if the webpage is the public face of the school and a surrogate measure of quality, it was still surprising that less than one-quarter of the medical schools had excellent and clearly understood webpages with readily navigable information on such characteristics as programmes, courses, location and accreditations. If, as discussed in Chapter 4, student recruitment, institutional reputation and the accessibility of information to stakeholders are the 'key drivers' of UK universities websites (Cox & Emmott, 2007; Schimmel et al, 2010), then the issue needs further exploration. It also begs the question of why three medical schools kept their FD provision private. A review of the existing literature indicates that while the educational benefits of learning via the web have been studied at length and there is a recognition of the importance of information and communications technology (ICT) for the future development of higher education institutions (HEIs), there is a paucity of data relating to the management of webpage content itself within HEIs (Cornford & Pollock, 2002). Other authors have also highlighted the diverse backgrounds of those in web management, their occupational

trajectories, responsibilities, job roles and the division between marketing and IT approaches to website design, purpose and maintenance (Cox & Emmott, 2007; Cox, 2007). This might be partly responsible for some of the issues noted such as lack of currency, coverage, comprehensiveness and poor functionality of the webpages as well as the disparity among schools.

In its investigation into the management of web content in HEIs in the UK, Eduserv (2009) concluded that the websites of HEIs in the UK have a multitude of functions and their audience is diverse. While a corporate site, for example, might attempt to cater for the needs of three or four different audiences, HEIs identified in excess of thirty target markets for their sites. University websites are used simultaneously to recruit new students (both domestic and overseas), to increase their exposure and to highlight their various achievements. They have to address the needs and requirements of the current student body and faculty while at the same time appealing to alumni, business partners, media and the wider research community. They also have to balance the demands of academia with commercial considerations that are increasingly becoming a more prominent part of university life. The challenge for website content was aptly summarised as follows,

> ***"For the content managers, understandably, this is somewhat of a 'headache' – particularly when one considers that web teams generally comprise a small number of people and are often situated (sometimes uncomfortably) between departments whose agendas may not necessarily be complementary." (Eduserv, 2009, p. 7)***

The Higher Education Funding Council for England (HEFCE), along with the Scottish Higher Education Funding Council (SHEFC) and the Higher Education Funding Council for Wales (HEFCW) are the funding councils whose remit it is to develop

policy and allocate public funds to the universities. There was suggestion by Eduserv (2009) that if there continues to be significant disparity between institutions' perceptions of the web's value and the resources willingly assigned to web development with poor end user satisfaction, then further investigation of the issue may be needed. The implication was that there is more that universities (and medical schools) could do to improve their webpages to make them more attractive, dynamic, interactive and user-friendly.

# CHAPTER 7

# PHASE II FINDINGS

## 7 Introduction

As described in Chapters 4 and 5, the purpose of phase II was to observe a particular FD programme in action and interview participants during and after the course to find out which mechanisms were activated and the outcomes that followed. This phase of the project was intricate in that it was carried out in steps to provide layers of data. The observation and short informal interviews during the course provided data on the three common mechanisms as explained in section 5.3. Six months later, the in-depth interviews of twelve participants provided further *CMO* data to support, modify or invalidate the hypotheses.

I start the chapter with a summary of the interview of the FD coordinator to put into context the background in which the course was held (Box 7.1). Next, I discuss the findings during the course followed by the participants' data from the six months follow-up. This is followed by a discussion of the views of the educators on how they see the future of FD in UK, then a brief discussion of the key findings from this phase and how they contributed in building the layers of this realist study. As mentioned before, a more detailed discussion of the findings from all phases is left till Chapter 9. I end the chapter with a summary of the important findings in this phase.

**Box 7.1 Summary of the FD Coordinator Interview**

The coordinator started in post 18 months prior to the interview in Jan 2012. She has a master's degree in computing and a PhD in education. She had worked in teaching and learning support in the NHS for over 10 years prior to the current post at WMS. Her job title is 'teaching and learning specialist' with responsibility for FD for clinicians who teach WMS medical students (a graduate entry course) as well as leading on the first module for the masters in medical education (the ECE course). The FD set-up is double-faceted. There is the WMS based FD office staffed by 2½ WTE (whole time equivalent) with two individuals focused on clinical educators in the acute NHS Trusts and a half time assistant post with particular responsibility for primary care and GPs. In addition, people directly employed by the university are able to have their FD training centrally through the university's learning and development centre (LDC). The source of funding for the FD office was an important issue as it relates to mission and governance. According to the coordinator, the university receives a small amount of service increment for teaching (SIFT) money from the LETB (local education and training board) and uses part of that for FD (approximately 1.5 WTE). Furthermore, the Postgraduate Certificate in Academic and Professional Practice (PCAPP) is free for university staff and the Postgraduate Certificate in Education (PGCE) is free for clinical educators who teach WMS students. The key objective of her role as stated in the job description was to design and deliver effective educational development opportunities for clinical teachers and develop FD activities across all clinical sites.

The coordinator highlighted that governance is a thorny issue in various aspects. There has been no systematic needs analysis of the WMS teachers. While it was mandatory for new university lecturers (on probation) to do the PCAPP or PGCE, there was no similar requirement for the clinical educators (they attend voluntarily). Attendance records were kept for courses held in-house but difficult to collect for those held outside the university. Similarly, some courses have a clear quality assurance system (e.g. the ECE course is accredited by the HEA as it aligns with UK professional standards framework) but others do not have a clear quality assurance system.

Focusing on the ECE course, she explained the key learning outcome was for participants to gain an understanding of how to deliver education in clinical practice as well as a basic knowledge of the underpinning theory. This aligns with WMS objective in delivering medical students education. For evaluation, she used three indicators. First was the standard evaluation questionnaire, which was the WMS form with a four point Likert scale covering four areas (teaching quality, module content, module virtual learning environment and overall opinion) and free text space; second was the thirty minute evaluation discussion with participants at the end of the course which she found very useful as the module leader to obtain feedback on how the module had gone. These two were processed along with her own reflections and comments from other facilitators to produce a report on each course. The third aspect was the assessment / grading of participants' reflective portfolios submitted three months post course by the facilitators. The argument was that since the portfolio tasks (Table 5.3) were aligned with the module learning outcomes, poor performance pattern in an area might flag up issues to be addressed in subsequent courses.

Regarding impact, so far she had been unable to demonstrate the impact of the FD for a number of reasons. One was the time frame as she had only been in post for 18 months, second was the difficulty with measuring educational change and third was the varied and dispersed settings from which participants came making follow up difficult. However, on further questioning, she admitted that the long-term impact she was hoping for is for FD to become institutionally visible and accepted i.e. teachers' development /supporting teachers in developing teaching skills becomes a routine part of the institutions activity and is embedded in the agenda of meetings, visible on websites, in publications and in departmental meetings. However, she readily admitted that this would not answer sceptic's questions about the evidence for impact, cost-effectiveness and improvement.

## 7.1 Demographic details of the cohorts

There were 22 participants in the January cohort and 24 in the April cohort. Sixteen participants out of 22 (73%) were interviewed during the January course and seventeen participants out of 24 (71%) were interviewed during the April course. The demographic data of the 33 participants are summarised in Table 7.1. The two groups had a similar demographic distribution in terms of gender, seniority, job type and funding for attending the course. However, there was a difference in previous teaching course attendance as only five participants (31%) in the January cohort had attended a previous teaching course compared with ten participants (59%) in the April cohort.

**Table 7.1: Demographic details of the 33 participants**

| | Jan Cohort | Apr Cohort | Total |
|---|---|---|---|
| **Gender** | | | |
| Male | 8 | 6 | 14 |
| Female | 8 | 11 | 19 |
| **Previous Teaching Course attendance** | | | |
| Yes | 5 | 10 | 15 |
| No | 11 | 7 | 18 |
| **Funding** | | | |
| Self | 6 | 6 | 12 |
| Others | 10 | 11 | 21 |
| ***Career Stage** | | | |
| Early | 4 | 3 | 7 |
| Mid | 6 | 6 | 12 |
| Established | 6 | 8 | 14 |

*Career Stage Description
Early: Training / less than 5 years post qualification
Mid: 5 – 10 years post qualification
Established: More than 10 years post qualification

## 7.2 Descriptive data - observed engagement scale

The observed engagement scale (where 1 represents no engagement and 5 represents full engagement) showed the mean observed engagement score for the January cohort to be 3.88, median was 3.7 (range 1 – 5). For the April cohort, the mean was 3.22, and median 3.4 (range 1 – 5). There was no statistically significant difference between the two groups. The median values showed that participants were ‘partly engaged’ to ‘mostly engaged’ during the sessions. The overall observed engagement score mean was 3.46 for the 33 participants. In addition, participants’ engagement scores were plotted for each session on slope diagrams to show the engagement trend over the three-day period. Slope diagrams are useful in giving a visual identification of outliers in terms of both absolute peaks and dip in engagement and to compare trajectories to see if less engaged people have different pattern from the more engaged. I studied plots of whole cohorts but these were crowded, and so here I have reproduced the trajectories of the least and most engaged participants from each cohort as well as a comparison of the course medians as discussed below.

Figure 7.1 shows the engagement scores of the least engaged participant (7J: 7 is participant code, J = January course) and the most engaged participant (18J) plotted against the median engagement of the January cohort over the three-day period. Similarly, Figure 7.2 shows the engagement scores of the least engaged participant (24A, A = April course) and the most engaged participant (16A) plotted against the median engagement of the April cohort. There were more fluctuations in the slope diagram of the least engaged April participant (24A) compared with the least

engaged January participant (7J). Participant 24A at least found some sessions quite engaging, especially session six ‘on the job teaching’ and session eight ‘teaching and learning in large groups’. The January participant (7J) was more consistently unengaged. A summary of the two unengaged participants (outliers) is included in Appendix 5.

**Figure 7.1: Slope Diagram of Engagement Scores Jan 2012 cohort**

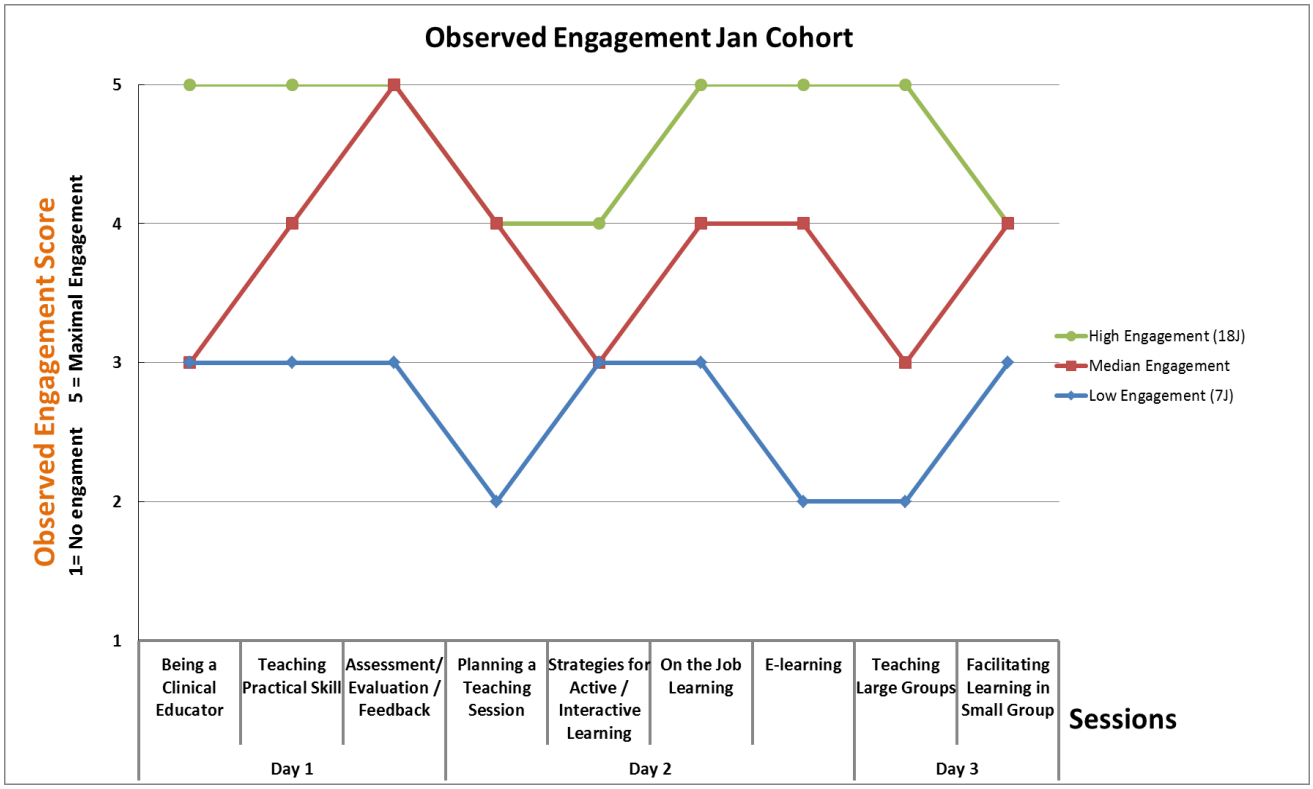
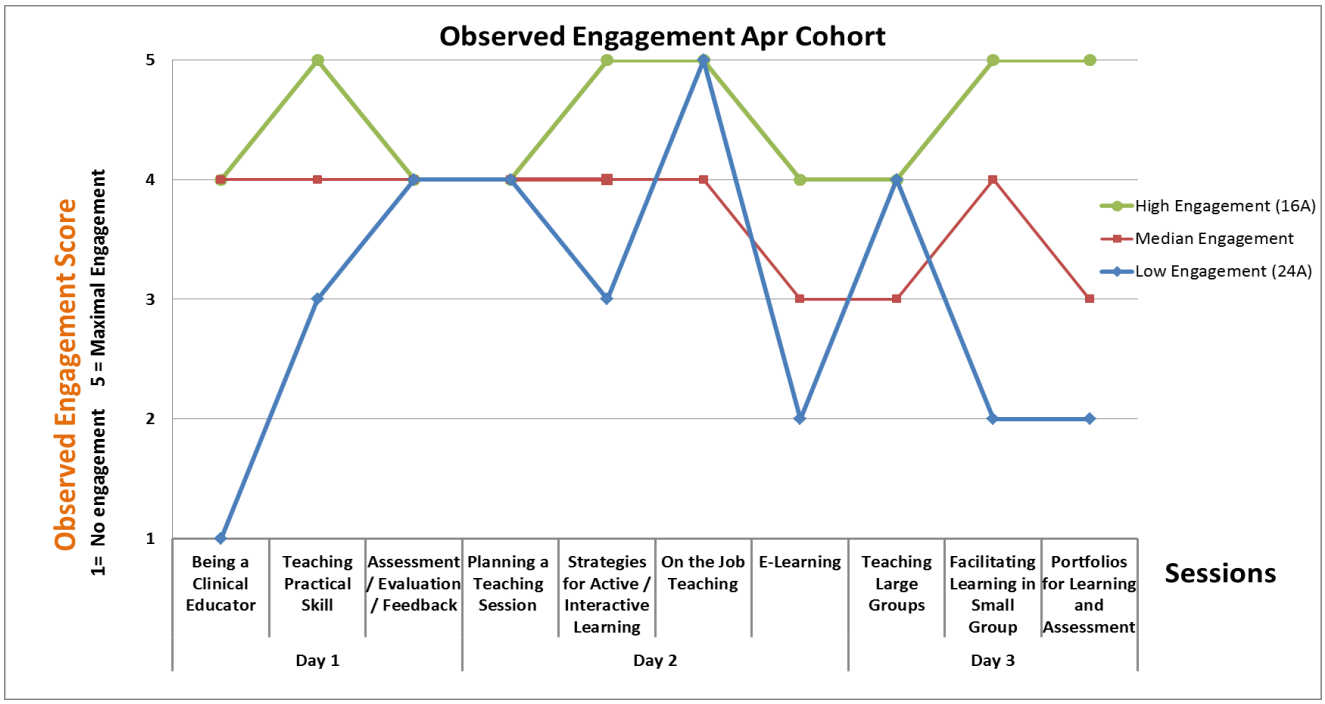
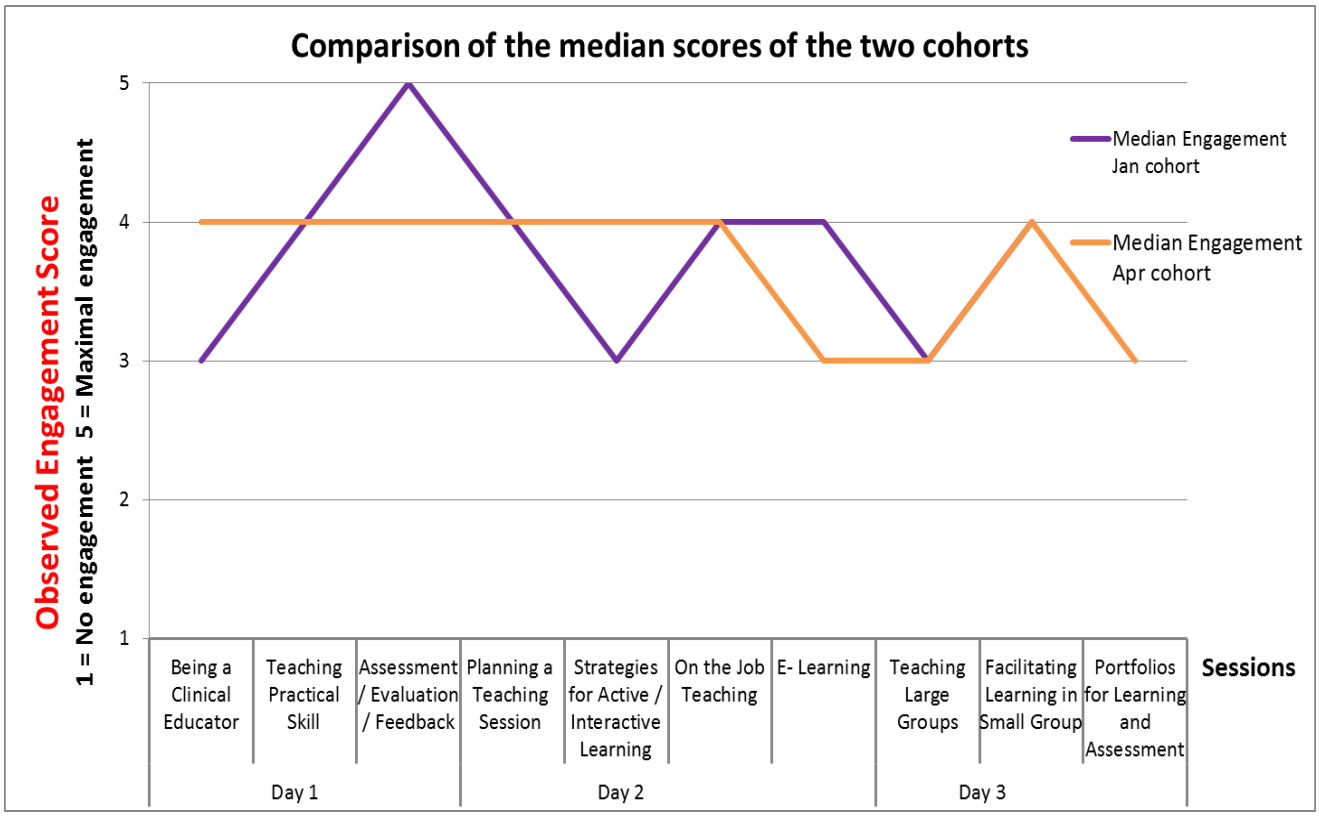


**Figure 7.2: Slope Diagram of Engagement Scores Apr 2012 cohort**




There was no statistically significant difference between the overall median engagements of the January participants compared with the April cohort as shown in the slope diagram of the medians in Figure 7.3. Admittedly, there was one outlying point in January with a score of five but this did not make a difference. This similarity of engagement mighty perhaps be explained by the similarity between the two cohorts and the fact that no changes were intentionally made to the April course in the light of January cohort feedback; I was present at both and this was also confirmed by the FD coordinator (see Box 7.1 paragraph four for the three ways that the FD coordinator gleans feedback used in altering future courses).

**Figure 7.3: Slope diagrams of the two cohorts median engagement scores**




## 7.3 Constructs data

To further explore the first three mechanisms (motivation, engagement and perception) listed under the hypotheses of the inquiry (Table 3.2) bi-axial constructs were developed from the course interview and observation data as discussed in section 5.3. The three mechanisms were all directly commented on by the 33 participants during their interviews. This helped in my development of the constructs using participants' coded data to derive categories which were distinct enough to be placed at polar ends of the axes. To aid understanding, I have provided working definitions of motivation, engagement and perception used in this thesis (Box 7.2).

**Box 7.2: Definitions of Motivation, Engagement and Perception**

(Martin, 2007; 2008; Reeve, 2001; Schaufeli et al, 2002; Bernstein, 2013)

Motivation is defined as a set of interrelated beliefs / emotions that influence and direct behaviour. Motivation is the impetus behind what a person actually does, the interior mental state that leads to action. It influences what people choose to do, how well, and for how long. Motivation is generally thought to be that which gives behaviour its energy and direction.

Engagement is defined as a positive and fulfilling learning-related state of mind that is characterised by vigour, dedication and absorption. Engagement is the link between what learners do, between the inner mental states of motivational and prosocial orientation and learning success. It involves a sustained thoughtful attention to learning with cognitive, behavioural and emotional components.

Perception (from the Latin perceptio, percipio) is defined as the organisation, identification, and interpretation of information in order to represent and understand the environment or context in which it happens. It is an intuitive recognition or appreciation of the qualities (moral, psychological, aesthetic) associated with information from the external world.

### 7.3.1 Motivation

Motivation was analysed on a bi-axial construct: *external* versus *internal* and *individualistic* versus *altruistic*. Two examples of participant responses categorised as *altruistic* are:

> *"Teaching is like performing, it's quite a good rush in a way. You get that warmth in you from helping the students, and if they like it, then it's a good feeling" (10A).*

> *"Teaching is something that I really enjoy that's why I do it. I think it is an important part of everyone's job. You have to do it well and actually one of the things about teaching is that it reminds you, about being the best doctor doesn't it really?" (17A).*

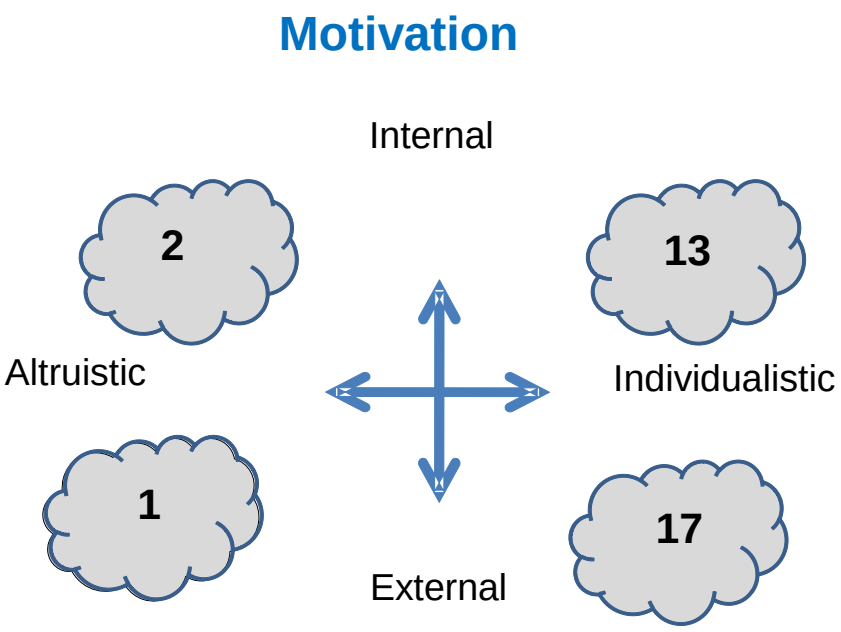
An example of participant response categorised as individualistic is:

> *"I want to equip myself now to become a better teacher in future. It looks good on my CV and will enhance career prospect" (20J)*

For the January cohort, the predominant motivation for attending the course was *individualistic*, i.e. related to personal need (n = 14 participants) rather than *altruistic*

n = 2). There was no difference between *external* (9) and *internal* (7) motivating factor. Similarly, for the April cohort, the motivation for attending the course was predominantly *individualistic* rather than *altruistic* (16 vs.1) again with no difference between *external* (9) versus *internal* (8) motivating factor. Thus in total, a statistically significant majority of 30 out of 33 participants ($p$ = 0.012) studied across both courses reported a predominantly *individualistic* motivation for attending. However, as Figure 7.4 shows, the external (18) and internal (15) motivating factors were similar ($p$ = 0.579). Further analysis showed no correlation between motivating factor, funding source, previous attendance on a teaching course or seniority status.

**Figure 7.4 Motivation: Individual motivation is the most common**




Examples of internal and external motivations are:

> *"I see learning how to teach properly a challenge for myself" (3J)*
> Internal (and individualistic)

*"I am interested and enthusiastic about teaching, I want to improve to help my students" (4A)* – Internal (and altruistic)

*"Attending this course will show my commitment to teaching as I'll put it on my CV" (8J)* – External (and individualistic)

*"I want to become a university lecturer hence the time and effort I'm investing in myself by attending this course" (13A)* – External (and individualistic)

*"To formalise what I have been doing for a long time and get a qualification in teaching. It will help me understand what is valuable to my students" (10A)* – External (and altruistic)

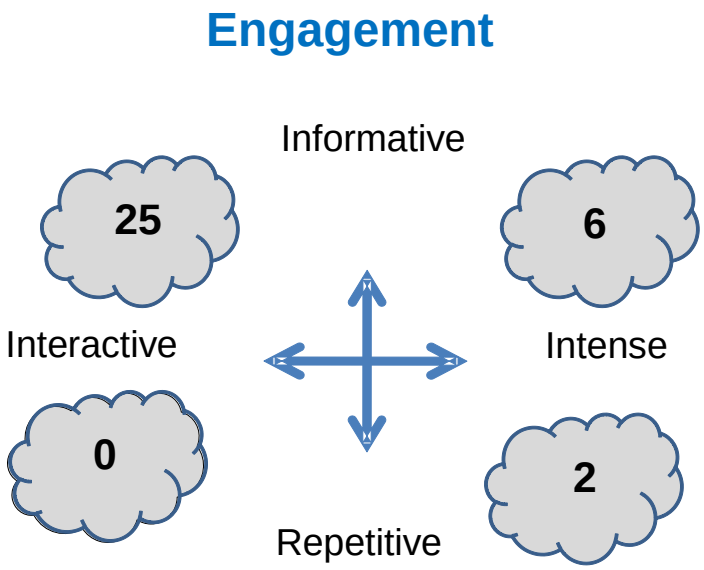
### 7.3.2 Engagement

For engagement, the bi-axial constructs were *informative* versus *repetitive* and *interactive* versus *intense*. In the January cohort, 12 participants were fully engaged in the *informative* / *interactive* quadrant, three participants were partially engaged (*informative / intense*) and one participant (7J) was unengaged reporting the course to be *repetitive*. In the April cohort, 13 participants were fully engaged in the *informative / interactive* quadrant, three partially engaged and one unengaged (24A) Overall, 31 participants found the sessions engaging (*informative* quadrants) and only two were unengaged (Figure 7.5) and this was statistically significant ($p$ = 0.011). Participatory learning utilising an interactive multimodal process was the most commonly reported contextual factor in maintaining engagement during the course.

As the researcher, it was interesting to relate my recorded observations of apparent engagement to participants' perceptions of their own engagement. There was good correlation between the observed engagement data and participants' perception of engagement as the two participants (7J and 24A) who appeared unengaged, had the

lowest observed engagement scores in their cohort and they reaffirmed later during their interviews that they found the FD activity *repetitive / intense*. I have provided some examples of participants' quotes reflecting the engagement construct below.

**Figure 7.5 Engagement: Engaged 31; Unengaged 2**




Informative (and interactive) examples:

> *"I found the sessions informative, interactive, engaging and kept me focused. Personally, observing alone is not good enough for me. I need to be doing something active to keep focused" (9J)*

> *"The sessions were very informative and have given me the science that underpins what I have been doing intuitively over the years. The interactivity was brilliant, for example, the practical skills sessions was very engaging" (12J)*

> *"A lot more interesting and informative than I expected it to be with a lot of interaction and participation. Very easy to give your opinion without intimidation; not judgmental, and open dialogue" (10A)*

Intense (and informative) example:

> *"I really have to pay attention as we have to feedback each time hence I find it difficult to relax. There was the tension of just waiting for the next exercise. I still found it engaging as it was informative but too intense for me" (3J)*

Repetitive (and intense) example:

> *"For me some sessions were quite repetitive, boring and too long, some were okay. I felt pressured from day one; I felt some things were skimmed over. I found the afternoon very intense and difficult to engage towards the end of the day" (7J)*

### 7.3.3 Perception

The bi-axial constructs for perception were *useful* versus *unproductive* and *relevant* versus *irrelevant*. Most participants (29: 13 from January and 16 from April) perceived the FD for teaching as *useful* / *relevant* on the basis that it gave them the confidence to practise various teaching methods and improved their teaching skills as noted in the quotes below.

> *"I have found the session very useful and relevant, it has been useful to get the structure of what I have been doing intuitively. It has given me more confidence to approach teaching. I have a better understanding of various teaching methods" (12A)*

> *"I found the sessions very useful, engaging and thought provoking. I think it will change my practice in giving feedback, designing feedback forms, and planning a teaching session. This is quite useful and relevant for me in a GP setting as I do one-to-one teaching" (19J)*

Two participants (one from each cohort) felt some of the sessions were partly *irrelevant* to them but overall still found some aspects of the FD useful. Interestingly, one participant was an experienced educator who has been teaching and running courses for years while the other participants was quite junior as only in the second

year of specialist training (ST2) and taught only on ad-hoc basis. Experience and seniority did not seem to make much difference in how they perceived the FD course as reflected in their quotes below.

> *"The afternoon session on assessment and feedback wasn't very relevant for me, it was boring and some of the sessions have not been very useful but a few have been useful" (7A)*

> *"I found some of the sessions too irrelevant and boring. Some were useful knowledge and skills if you apply it right away (short-term application)" (13J)*

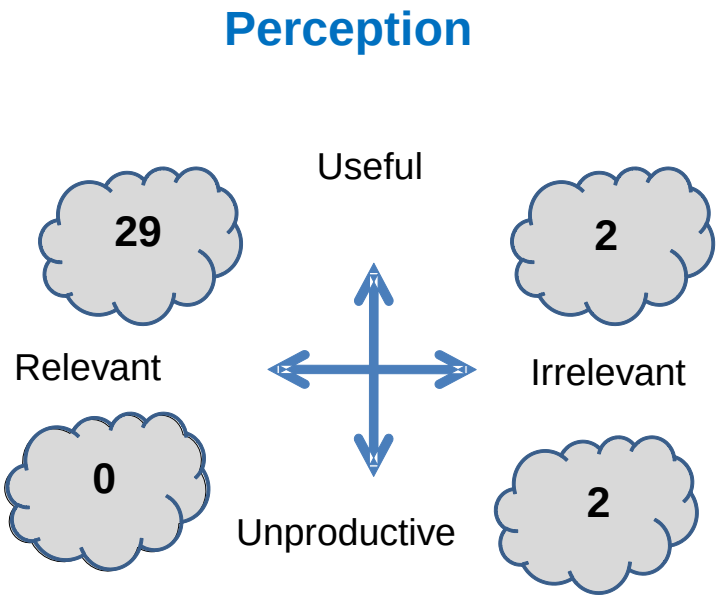
Only two participants (one from each cohort) found the course both *irrelevant* and *unproductive.* In Figure 7.6, I have summarised the perception construct showing 31 participants found the FD activity *useful* and only two found it *unproductive* ($p$ = 0.011) which was statistically significant. There was good correlation between the data sources (the observed engagement scale and the constructs derived from the interview data) as the two participants describing the session as *irrelevant* / *unproductive* were the two already identified above as appearing disengaged to the researcher. Apart from gender, the two participants (7J and 24A) didn't seem to have much in common. One was self-funded, the other funded by his/her organisation, one was established career the other mid-career, one in education the other in clinical practice, one more consistently unengaged and the other more fluctuating. Interestingly at the six months interview, participant 7J who was consistently the most unengaged, felt on reflection he/she was probably a bit too judgemental because having already done the PGCE, he/she found most of the sessions repetitive but still learnt few useful tips (Appendix 5). As the two participants were consistent outliers, it was important to consider them in more detail hence I have

summarised their details in Appendix 5 but I have provided below, some extracts of their responses during the course interviews.

> *"Most of the sessions have not been useful or relevant for me. I wonder why the course is for 3 days and not 5 days as the others especially as it costs the same. There was no chance to delve into things in-depth" (7J)*

> *"Some of the sessions were irrelevant and lasted an hour or longer when probably they could be shorter. I found the small group session too long, it could be shorter. I am not sure I gained much in terms of productivity" (24A)*

**Figure 7.6 Perception: Most participants found FD useful and relevant**




## 7.4 CMO Data

All the coded data in phase II were categorised under Context (C), Mechanism (M) and Outcome (O) in line with the *CMO* hypothesis development of the realist framework (Pawson, 2013). In the sections below, I have detailed the findings under

each subheading with the supporting data added as well as highlighted the relevant hypothesis of the inquiry.

### 7.4.1 Contexts

Context refers to those features of the conditions in which interventions are introduced that are relevant to the operation of the programme mechanism, i.e. it facilitates the effectiveness of the programme. I identified possible contextual factors during and after the course (Table 7.2) and explored their influence on the mechanism and outcomes.

**Table 7.2: Contextual Factors**

| | |
|---|---|
| Previous teaching course | Participatory approach |
| Lack of pedagogy | Reflective practice |
| Learning with / from peers | |

#### 7.4.1.1 Previous teaching course

In terms of previous attendance at a teaching course, five participants in the January cohort (two of whom had a PGCE) and ten participants in the April cohort had attended a previous teaching course i.e. 15 out of 33 participants. The types of courses attended were quite varied and the time passed since those courses ranged from one to eleven years with the majority being more than five years before the ECE course. Apart from the two participants (7J and 12J) with PGCE, the other courses focused more on practical tips or addressed a specific area such as assessment and feedback. The expectation would be that prior attendance on a

teaching course would influence participants' in some way. However, except for 7J (already discussed above), there was no difference between participants with prior teaching course attendance and those without in terms of their observed and perceived engagement, their motivation or their perceptions of the ECE course. This might be due to the type and nature of the previous courses, the time expired since the course, participants' need for reinforcement or the fact that prior attendance does not mean that learning has taken place. So, previous teaching course attendance doesn't seem to be activating any of the mechanisms under consideration.

#### 7.4.1.2 Lack of pedagogy

About half of the 33 participants from both cohorts mentioned during their interviews, lack of knowledge about education and learning theories prior to the ECE course. Interestingly, apart from the two with PGCE, there was no correlation with previous teaching course attendance. This again may reflect the fact that attendance does not equate to learning or the types of courses attended were not focused on the theoretical aspect of education. The ECE course made them aware of the theoretical framework underpinning teaching approaches. There were at least 17 references to lack of pedagogy (both from people that had attended other teaching courses and those that hadn't) signifying the importance of this issue to participants. In my contemporaneous observation notes, I have commented that participants were glad to have this knowledge to back up their teaching i.e. what they have been doing intuitively. In their group discussion, they commented on how it was now easier to explain why they would use a particular teaching method. However there was no clear separation between those that discussed theoretical knowledge and others that

didn't with regards to their overall perception of the course or how engaged (observed and self-reported) they were during the course as noted in my observer records. Below I have given some examples of participants' extracts illustrating lack of pedagogical knowledge prior to the course.

> *"Teaching in the university I don't have any pedagogical knowledge. Personally I am relieved to have some formal training in education, to learn these theories and different styles because I've never had any training before" (18J).*

> *"Prior to the course, one of the big gaps was educational theory because I wasn't aware of the literature that was out there and the broad principles. Now I am better grounded in the theoretical side of medical education… it (the FD) gave me a framework for my teaching. It has given me the underlying principles and theory where there was a big gap before" (6J).*

Lack of pedagogy and subsequent understanding of theoretical knowledge is probably a contextual factor in so far as participants were glad the course addressed the gap in their knowledge rather than in terms of activating a mechanism. However, one could argue that for a small subset of participants FD aligned with, or relevant to, the needs of the participants would motivate them to learn better and improve their instructional knowledge. The hypothesis supported by is hypothesis 2 (Table 7.3)

**Table 7.3: Hypothesis related to lack of pedagogy**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 2 | FD course aligned with or relevant to the learners' needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills and change in teaching approach |

#### 7.4.1.3 Learning with and from peers

The two cohorts were small groups of highly educated individuals who were able to share teaching experiences, problems and solutions. Over the three days of the course (8 hours daily), the group learned together and from each other but the rest of the programme involved 175 hours of self-directed learning. Participants seemed to enjoy learning with and from each other as evident by some of their quotes below.

> *"When I learn the most is when we discuss things. It is good having the handouts and the slides but it's when we discuss things, that's what I enjoy most" (23A)*

> *"Many things can be gained by talking to like-minded people, hence the importance of this course. The environment helps to stimulate that and I gained a lot" (3A)*

Peer learning as explained by Boud (2001) essentially refers to,

> ***"students learning with and from each other as fellow learners without any implied authority to any individual, based on the tenet that students learn a great deal by explaining their ideas to others and by participating in activities in which they can learn from their peers" (Boud, 2001, p. 4).***

Boud (2001) further went on to explain that learning with and from each other is a necessary and important aspect of all courses and, while the role it plays varies widely and the forms it takes are very diverse, without it students gain an impoverished education. He suggested that peer learners would engage themselves *intellectually, emotionally* and *socially* in *constructive conversation* and learn by talking and questioning each other's views. Peer learning can be further enhanced if the environment of mutual help continues over time and beyond the classroom as learners become individually and collectively accountable for optimising their own learning and achievements (Boud, 2001). Technology is now an important driver

towards the use of peer learning and web based activities appear to be most effective when there is interaction among the learners themselves (Boud, 2001; Stephenson, 2001). Other authors have expressed similar views: Bohuijs (1998) commented that information technology has provided students with excellent opportunities *to learn with and from each other.*

However, in this study learning with and from peers was limited to the 24 hours of face-to-face contact because, despite having an online forum set up specifically for each cohort, it was hardly utilised to generate discussion, ask questions, pose problems, express difficulties encountered or suggest solutions. This might not be uncommon but the fact remains that virtually all the participants interviewed post course expressed the challenges, difficulty and sense of loneliness that they experienced in doing the post course assessment, but did not use the online discussion forum or contact their peers. This point was eloquently expressed by one of the participants,

> *"I'm not very confident or comfortable discussing with colleagues through that medium (online) as it's not something that I'm very used to, so that might be why I didn't engage with it" (6J).*

There were various other reasons put forward by participants including time pressure, competing demands, lack of familiarity, lack of awareness of the scope of its potential use and lack of technological know-how. Even though they seemed to have enjoyed the discussion and learning from each other during the course (as this was the way the course was structured), once released from this, the few groups that kept together were all at the same career level. This might suggest that the hierarchical nature of medicine could also have a part to play in addition to educators

not wanting to venture out of their comfort zone i.e. technological know-how and generation differences in terms of using social media.

There is no doubt that peer learning could be a context that leads to good educational outcomes as learners who engage in effective peer learning perform better academically, persist longer, feel better about the educational experience, and have enhanced self-esteem (Landis, 2000). However, in my view, while learning with and from peers was a useful strategy during the course, there was limited utilisation of peer learning by participants afterwards. I do not believe that the process was sustained enough for this to be a key influential context for the majority of participants (based on their comments) compared with other contexts discussed below.

#### 7.4.1.4 Participatory approach

During the course, all the sessions involved participants taking an active role in the sessions through group discussions on the tables, writing on flip charts and presenting the findings of the group. As noted in my contemporaneous observer notes, 50% of the time the groups were involved in one form of activity or the other and contributed to the learning points during the sessions. Participants described an interactive, experiential approach involving actual practice and learning from the activities as key to their engagement and learning during the sessions and in consolidating their knowledge and teaching skills. The majority of the participants found this participatory approach highly engaging; it was the most commonly reported contextual factor in maintaining participants' engagement during the course. Below are some extracts.

*"So I enjoyed the sessions as they were short and participatory. I would have found the sessions very dry, unengaging and quite unhelpful if they were totally lecture based and the delivery was not participatory" (3A)*

*"I found the mixture of sessions, group discussions and variety of activities was good otherwise I would have been more exhausted if it was just didactic. The variety of teaching methods / presentation was very engaging and one of the most important part for my learning" (21A)*

*"I found the participatory approach quite interesting and fascinating. The mix of activities made it very interactive and it made even quiet people like me talk and interact. I learnt a lot this way" (15J)*

Evidence from the literature supports a participatory approach using multimodal methods as being important in active learning. The underlying principle in participatory learning is to engage the participation of people in the processes of learning about their needs and opportunities. The key is the promotion of interactive learning, shared knowledge, and a flexible approach (Pretty et al, 1995; Chambers, 2013). Steinert's et al's. (2006) systematic review summarised effective FD as that using multiple instructional methods, experiential and participatory approaches as discussed in section 2.6.3. However, I was also cognisant of the fact that while the majority of participants found the active participatory approach very engaging and stimulating, two participants found it tiring and intense as discussed in section 7.3.2 above. My view is that, the participatory approach was a very important contextual factor for most participants as it stimulated and focused their attention. The context supports hypothesis 1 as shown below in Table 7.4

**Table 7.4: Hypothesis supported by participatory approach**

| | Context | Mechanism | Outcome |
|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + Engagement generated by the facilitator being very interactive and informative using multimodal approach | = Increased confidence in teaching. Empowerment to utilise teaching opportunities |

#### 7.4.1.5 Reflective practice

The data from all twelve participants interviewed six months after the ECE course showed that reflective practice was the key contextual factor that maintained their focus on teaching and learning. They commented that initially this was through the post-course assessment (reflective portfolio, teaching practice and peer observation) with action plans developed by the participants that were relevant to their own teaching, and later reflection was fostered through experiential practice in their own working environment. This view was strongly expressed by all twelve participants as the single most important factor that kept them learning beyond the course. Further data analysis revealed that eleven of the twelve interviewees commented on teaching observations with reflection being the most influential reflective practice they undertook. While there was no doubt that the sum of the various reflective activities they practised was important, it was still useful to know that the experiential component remained a large part of this process. Below are some extracts from participants on reflective practice.

> *"But really the big thing that changed my practice was the reflection during the assessment. So going away spending time … actually reflecting, putting something in practice into my own teaching and thinking about it, was the thing that's really changed me… so if one thing has changed my practice it was the reflective process of assessment" (7A).*

*"Going through my personal reflection side of it was important to consolidate and further my knowledge. In terms of knowledge through attending the course alone, I gained, but even more so through the write up, the directed reading and the reflective practice. The reflective practice helped me to gain even more as an educator….. The personal reflection is where the individual development and direction takes place as you are forced to focus individually on your own teaching environment and apply it to your own role. Without the reflective practice the application of my knowledge would not be as thorough. Kind of shape your teaching skills through your own experience and understanding" (6J).*

Even participants, who initially found reflection a difficult concept to understand or accept, later became advocates of reflective practice and eulogised about its importance to their learning.

*"The issue of reflection you know, it wasn't that I was not interested in knowing about it but it is something that initially was being forced into me but I have come to the understanding that it is important to sit back and think, how did it go? What else can be done? So I know now that it is important. To consolidate my understanding, it was useful to do the reflective practice. It really helped my learning because I had to think how I could do it better, plan something myself and put it into practice" (21A).*

To summarise, the data suggests that reflective practice was an important contextual factor and this fits with hypothesis 1 as shown in Table 7.5 below.

**Table 7.5: Hypothesis supported by reflective practice**

| | Context | Mechanism | Outcome |
|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + Engagement generated by the facilitator being very interactive and informative using multimodal approach | = Increased confidence in teaching. Empowerment to utilise teaching opportunities |

### 7.4.2 Mechanisms

As explained in Chapter 3, mechanism refers to that aspect of the programme which makes it work. As with any complex intervention, in FD activities, there are various mechanisms to be considered. The data were examined for findings to support, modify or invalidate the mechanisms proposed in the hypothesis of the inquiry (Table 3.2). The four mechanisms identified from the data are discussed below.

#### 7.4.2.1 Motivation

The predominant motivation was individualistic, i.e. related to personal need rather than altruistic as already mentioned under constructs in section 7.3.1. Four participants (two from each cohort 6J & 18J; 12A & 21A) were attending the course on a mandatory basis because of their contractual requirement (two lecturers and two GP clinical tutors); the others were all attending voluntarily. Most participants had at least two motivations for participating in the ECE. These could be viewed as a spread from personal oriented domain to teaching oriented domain with further subdivisions (or categories) as shown in Table 7.6.

**Table 7.6: Motivating factors**

| *Personal Oriented* | *Teaching Oriented* |
|---|---|
| Career development (17) | Personal teaching interest (12) |
| Personal development (10) | Specific instructional skills (8) |
| Peer influence (6) | Theoretical knowledge (6) |
| Optional module (5) | Professionalism (6) |
| (The numbers in bracket represents the number of times the motivating factor was identified by the participants) | |

The top three participants motivations were career development (improving CV, qualification, getting on in the NHS), personal teaching interest (enjoy teaching,

interest in teaching, opportunity to update) and personal development (become more confident, relief in having some training and validation to check whether what they are doing is right). This last point was neatly summed up by one participant,

> *"I don't know if I am teaching correctly, I enjoy teaching, but is my teaching effective? That is a completely different kettle of fish. Am I doing it the way I'm supposed to?" (24A).*

Interestingly, there were some people influenced by friends, peers or colleagues who had already done the ECE course, a PGCE or a Master's degree. A few were doing it as an optional module for a Master's degree. Some participants were able to identify gaps where they needed specific training such as evaluation, feedback and theoretical framework. Professionalism though, in terms of awareness of personal responsibilities as a teacher, role model, was mentioned by fewer than 1 in 5 participants. I also explored the issue of funding and how it might have influenced motivation to learn as approximately one third were self-funded and the other two-thirds were funded from various sources. There was no difference in the motivation pattern in the two groups. In fact one participant commented on this and said:

> *"You know the motivation would have been the same if I had paid for it myself or as in this case my NHS Trust had funded it. I feel an enormous obligation to the Trust to achieve but I would have felt the same had I funded it myself because you know if you are paying £900 then obviously you must achieve" (14A)*

Taken together, career and personal development as well as personal teaching interest were clear motivators for participants' choice and decision to attend the ECE course. Below I have given examples of individual motivation extracted from participants' data.

*"I am doing this for future career development. I work in a training practice and do one-to-one teaching. I enjoy getting involved and there is opportunity to do more teaching if I want. It is a requirement that GP trainers need a postgraduate certificate in education" (17J)*

*"I enjoy teaching, it is fulfilling to see someone demonstrate that they understand what you taught them, apply the skill and contribute to the body of knowledge in general" (15A)*

*"I am attending to improve my skills, make sure I'm doing the right thing and to become more professional in teaching. It will also improve my CV and career prospects. I'm doing this out of my own desire; it is not mandatory. I am doing more regular teaching, so need to be more confident" (18A)*

However, once the course got underway some of these individual factors became less important for some educators as participation was no longer dependent on them as noted in my observer comments. In my view, motivation was important in the decision making for attendance and for some participants it was the driver, as they had specific instructional areas (for example interactive skills and on the job teaching skills) that they were keen to learn about. Hence at least for a subset of participants, motivation was a mechanism and supported hypothesis 2 (Table 7.7).

**Table 7.7: Hypothesis supported by motivation**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 2 | FD course aligned with or relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills and change in teaching approach |

#### 7.4.2.2 Engagement

There were many factors that contributed to the course being highly engaging and stimulating. The three main ones that were readily identifiable are listed in Table 7.8. The multimodal approach including mini-lecture, role-play, participant presentations

and small group discussions was commented on by many participants as very engaging. In my observer comments, I documented that participants were involved in a lot of interactive task-based activities, encouraged to write on flip charts, practical demonstrations, which they found interesting and enjoyable. They readily reported this as a key factor in maintaining engagement throughout the course. Towards the end of each day there was time spent in a reflective period to check whether the learning outcomes had been achieved using metacognitive skills such as questioning and summarising. Participants interviewed during and after the course further identified engagement as a very important factor for their learning during the course.

**Table 7.8: Engagement Factors**

| | |
|---|---|
| Multimodal approach | Variety of teaching methods used to appeal to different learners (visual, auditory, kinaesthetic) |
| Interactivity | Highly interactive task based approach |
| Metacognition | Reflective approach making participants think |

*Examples of engagement during the course*

> *"I found the sessions well organised and planned. There was a lot of questioning from the facilitators, which forced us to think. I liked the interactivity, I found it more engaging. The timing was great as well; I find prolonged lectures are boring" (2J)*

> *"It was interesting and interactive, day two was more fascinating because there were things I never thought of before, they were eye openers. For example 'on the job teaching' – I can use situations that I will formerly have seen as a set back as a teaching opportunity e.g. if a patient did not attend, I can use the time to teach. It has been quite stimulating and I learnt a lot" (15A)*

After the course, participants still commented on engagement as being a very important mechanism for their learning during the self-directed period. The variety of the mandatory reflective portfolio tasks (seven in total – see Table 5.3) and their

shortness appealed to many participants compared with having one long dissertation for the assessment. One could argue that the mandatory tasks were the motivating factor but the educators genuinely wanted to learn more and commented that the experiential practice, reflection, teaching observations and feedback kept them engaged and learning during the post course period. This was pithily summed up by two participants,

> *"So the good thing was that the assessments weren't too long and were separated into small bits ….. obviously the aim was to get an understanding and to reflect on our own practice and to learn and shape them through our own experience and understanding by thinking about it … So this was helpful, it was specific, focused and very good for me to learn further" (18A).*
>
> *"I found at some points learning could be very challenging. You reflect on your own, on how you learn in your own way, in many ways it was a self-analysis. It was identifying how I learnt, why I learnt like that and the time spent after the course doing the reflective portfolio developed my learning greatly" (14J).*

From the available evidence, engagement was a very important mechanism that made the FD programme effective for participants. During and after the course, participants' engagement was reported as a key factor that led to learning. This supported hypothesis 1 as shown in Table 7.9 below.

**Table 7.9: Hypothesis supported by engagement**

| | Context | Mechanism | Outcome |
|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + Engagement generated by the facilitator being very interactive and informative using multimodal approach | = Increased confidence in teaching. Empowerment to utilise teaching opportunities |

#### 7.4.2.3 Perception

Perception as defined in Box 7.2 is an intuitive recognition, identification, interpretation and appreciation of the qualities associated with information from the external world. It was important to explore and interpret participants' perception of the FD course in terms of usefulness, relevance, applicability and overall experience (Table 7.10). The majority of participants had a positive perception of the FD and commented on the value of attending the course, the changes they have been able to make in their teaching (more interactivity with their students, clarification of objectives) and improvement in their instructional skills as evidenced by the peer observation of their teaching. Some went further to describe the experience as being very enjoyable and stimulating.

**Table 7.10: Perception Factors**

| | |
|---|---|
| Usefulness | The quality or capacity of having utility, being beneficial and / or having practical worth. A measure of how the participants saw the benefits and value of the FD they had attended. |
| Relevance | In teaching, a concept or information is deemed relevant if it increases the likelihood of accomplishing the goal e.g. in FD the training should be pertinent, important and connected to the teaching that the educators do. |
| Applicability | The fact or state of being suitable, appropriate or fitting to the issue under consideration. In this case, it meant the educators could try out the various skills and strategies in their own teaching. |
| Overall experience | Comprises the wisdom, knowledge or skill gained through involvement or exposure to an event (in this case the FD). It is the totality of the cognitions from that event; all that was perceived, understood, and remembered from the course. |

I have given some examples under perception constructs in section 7.2.3 above but I have included some participants' quotes below to further illustrate the issue of usefulness, relevance and positive experiences.

*Examples of usefulness / relevance*

> *"Fantastic, the FD training has been very useful. It's kept me up to date. There were many ideas that I have been able to use and it has made me a better teacher. It's made me change my practice in a practical way and made me question what I do. Previously I would have just got a piece of paper and write but now I think more deeply about the objectives of the teaching and why I'm doing it, so a much bigger focus on the objectives" (2A).*

> *"I found the sessions quite useful and relevant. For example the graph that showed how the attention span of learners decrease over a period of time and how to bring back their attention and productivity was very informative. I apply the technique now in my teaching and it has improved my teaching" (5J).*

*Examples of positive experiences*

> *"I really enjoyed it a lot more than I expected I would. I was interested in the education side of things but got an awful lot more out of it. I didn't think… well I thought I'd enjoy it but I really enjoyed it, and it started me thinking about teaching as a career" (16A).*

> *"It's been absolutely valuable for the work that I do as an educator. The insight that I now have in education has increased. It was an enlightening and enjoyable experience. It has influenced my teaching positively so I have a positive attitude and I'm glad I went on it. It's been useful; really invaluable experience" (14J).*

However despite the majority of participants describing a positive perception and good experiences, there were two individuals with very negative experiences, which were strongly expressed as shown in the extracts below.

> *"Just disappointing. Some of the presentations were disappointing and they overlapped as well, so you've got one presentation and then the day after you got very similar slides from the day before. It's like coordinating; making sure that the information is different for each lecture to avoid repetition" (20A).*

> *"I mean it's made me aware of theory and principles alright and from that point of view I think it's been useful. But as I say I don't think it's kind of significantly altered my clinical or teaching practice. I would*

*say that it is possibly better if it was more focused for senior clinicians" (4A).*

Despite this, the two individuals (4A and 20A) still participated actively as noted in my observer record. Their observed engagement scores were above the median suggesting that they were not in a negative mode during the majority of the course but on reflection six months later have expressed these negative comments more strongly. Another possible explanation for their active participation might be the small group effect. As half of the time was spent doing tasks in groups of four or five which then had to be presented or written on a flipchart, it was more difficult not to participate compared with just being part of a larger group of 24. The level of knowledge might also be a factor. The two individuals were both senior in their career with significant experience of medical education theory and practice hence the comments about the level at which the course was pitched. So while in my observer comment they actively participated and appeared engaged, probably no further deep learning took place presumably because there was no new information to learn.

I also explored perception by asking the participants who they felt was the client on the FD course, who was the focus on – the participants, the participants students or the organisation delivering the teaching? Eight of the twelve interviewees strongly felt they were the focus of attention. Of the remaining four, three felt their students were the focus of attention and only one felt the university had its own agenda as the focus. I have used two contrasting quotes below to illustrate the point but my overall impression was that the majority had a positive feeling that the coordinator and facilitators were focused on them and their needs over the three day period.

*"Personally I felt it was quite focussed on me as an individual because you can make it personal by taking things that people are saying in a fairly general way and make it more relevant to yourself, make it more focussed on yourself by reflecting a bit on how is that going to affect me?" (16J)*

*"You're tapping into my cynical side now because my cynical side would say that the university was only interested in the students getting the best teaching they can, they are not interested in the bit in the middle. There was quite a big focus on what students want, what students like, giving feedback and so on. But I don't think the ultimate aim was to develop the teachers per se. It's so they could say to the students as consumers … we have well trained teachers in our schools" (10A)*

On balance, the evidence was that most participants had a positive perception and a positive experience of FD. Therefore in my view, positive perception was a mechanism and supported hypothesis 3 as shown in Table 7.11 below.

**Table 7.11: Hypothesis supported by positive perception**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 3 | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Increased enthusiasm, interest and higher importance attached to teaching |

#### 7.4.2.4 Feedback

There was a variety of sources of feedback for the participants. During the course there was feedback from peers (during skill demonstrations, presentation sessions), and from facilitators (to the small groups and individuals at times during tasks). Post-course, there was feedback from the peer observation of teaching practice and from the assessors following submission of the portfolio. Most of the feedback was generic (rather than individualised) and based on a framework of giving constructive and meaningful feedback to enhance learning. Participants valued the feedback as it

gave them a framework to practice with and improved their understanding of the importance of feedback. Below are the extracts from two participants on how valuable they found the feedback they received from their peer observation of teaching.

> *"I did observe a colleague and my colleague observed me teaching. The feedback was useful; I got a lot of feedback which was very helpful in improving my teaching" (18J).*

> *"I have had feedback before but everyone is very praising but it's more the constructive criticism that was really useful. I did my peer observation teaching with a girl who has done the course before, so she knew exactly what I needed… what to look out for and how to feedback to me and vice versa so I got quite an awful lot from that. To get structured feedback and to use peers that are interested in education as well to observe you was very useful" (16A).*

Thus from the above, feedback to some degree could be conceived as a possible mechanism leading to improved teaching outcome. However, in my observer notes and the interview records, I noted that feedback was one of the specific skills that some participants were deficient in and the course did provide them with a strategy to improve their feedback skills, utilise and practise feedback having seen examples of how it was done. Some participants gave examples of how they have altered their feedback in various scenarios such as in marking written papers, practical skill demonstration, simulations and clinical teaching as shown in the two quotes below. Viewed this way, deficiency in feedback skills could be considered a possible context, which motivated learners to learn and improve their feedback / instructional skill which supports hypothesis 2 (Table 7.7) above.

> *"Yes, certainly there were gaps in my knowledge and practice and one was certainly this issue of giving feedback which I have worked on. I've come to appreciate how important it is and I now know how*

*to provide it. It is much easier for me to do and knowing that it is important, I am more informative than I used to be. For example in students written assessments, I barely used to write two sentences as comments but now they get more feedback from me" (9J).*

*"I certainly focus more on feedback … I try to be quite positive in my feedback, so I've changed the way I give feedback, I've tried to make it more constructive rather than just a criticism" (7A).*

On balance, feedback was at least a possible mechanism for some participants even though it could be argued that it was not as strong as some of the other mechanisms already considered. However, deficient feedback skill was also a possible context for some participants. This issue of allocating factors to context or mechanism has already been discussed in section 3.4. I will explore feedback in more detail under the discussion of all findings in Chapter 9. For now, there is at least some evidence to support feedback as a plausible mechanism for some participants as shown in hypothesis 4 (Table 7.12).

**Table 7.12: Hypothesis supported by feedback**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 4 | FD using an iterative cycle of training, changes to course design and continuous dialogue with stakeholders | + | Feedback to the educators during the FD and / or assignments | = | Improved teaching performance |

### 7.4.3 Outcomes

Participants when interviewed about the effect of the ECE course six months later mentioned various outcomes. The outcomes can be broadly grouped into two domains: teacher development and personal development as shown in Table 7.13 below.

**Table 7.13: Outcomes**

| *Teacher Development* | *Personal Development* |
|---|---|
| Instructional skills | Confidence |
| Teaching methods | Empowerment |

#### 7.4.3.1 Teacher Development

Participants gave many instances where they have been able to utilise, incorporate and implement the skills learnt on the course. Interactive skills such as fish bowl technique, line-up technique were mentioned as examples. They used these skills to engage their learners in active learning rather than passive transmission of knowledge, which tends to lead to 'inert' knowledge i.e. knowledge which might be acquired by the learners but not utilised (Bransford et al, 2000). Other skills utilised were in the areas of feedback, evaluation and assessment. Teaching methods were another aspect highlighted by the educators as an area of change. Small and large group teaching, on the job teaching, mini-lectures, and presentation by learners were examples of the various methods they have successfully used. Even though they were aware of some of these before, the course helped to improve their content delivery using these methods. Some examples of teacher development are given below.

> *"I really learnt a lot about education theory. I learned a lot about how to teach students using the different modalities that are there. You know, it gave me experience about how to give a lecture, teach a small group, set an exam, and how to organise a module of a course even" (11A).*

> *"I think the course has been very useful and productive for me as I've not done such interactive teaching before but actually I'm preferring it now. It is harder work in a way as you feel less in control (I think I'm probably quite controlling) but I'm learning to let the control of the session swing more towards the students" (18A).*

#### 7.4.3.2 Personal Development

Interestingly, it is in the domain of personal development that the educators noticed the greatest impact. They felt empowered to recognise, manage and utilise teaching opportunities. Prior to the course, they were hesitant to take on teaching challenges or opportunities, but following the course, there was a definite shift in their approach. Examples they gave include delivering a lecture on behalf of a senior colleague, starting a journal club, representing GP registrars (trainees) on the Trainees Association of Medical Education (TASME), planning student experiences during specialty rotation and organising multidisciplinary teaching. Educators now feel they have the credibility to manage or utilise these teaching opportunities where there was doubt before. This was an important shift.

The outcome which attracted most participants' comment was the increased level of confidence in teaching, which was expressed by all twelve participants at the six month's interview. Confidence in this context was defined as the ability to be certain and was mentioned in relation to three key areas: delivery methods, design / planning of teaching sessions and discussion with colleagues about teaching. When I explored further, I found confidence to be independent of the seniority of the participants, which was perhaps a bit surprising as one might expect the more established educators (e.g. consultants) to already have this ability compared with the early stage career educators. A possible explanation for this was that the confidence wasn't to do with content knowledge; rather it was confidence in the other aspects of the teaching process. In my observer notes, I did comment that (with very few exceptions) in both cohorts, there seemed to be a degree of doubt and

uncertainty on the part of the educators regarding the design and delivery of their teaching. Some expressed this as experiencing a feeling of uncharted areas regarding the principles underpinning the learning process, hence the confidence later developed following the FD programme. Two participants' quotes about their increased confidence in teaching are given below followed by quotes about feeling empowered.

*Examples of confidence*

> *"Yes I do feel it's improved my confidence. I feel more conviction that I'm going in the right direction now because previously I had to be guided by gut reaction whereas now I feel I have actually got some evidence to back up what I am doing. It makes me feel more confident that I am on the right track" (6J).*

> *"As an educator, I feel a lot more confident in myself and not necessarily with the content of what I am teaching because my knowledge of that has not changed. To be confident as a teacher does matter you know; the students don't find it easy to learn from a teacher who doesn't appear to be sure" (10A).*

*Examples of empowerment*

> *"The thing that I do more now than before (knowing that I was more a transmission teacher), is I actually feel more empowered as an educator to let learners find things out themselves (if appropriate). I feel much more empowered and confident about doing that now as I'm aware there are lots of different ways that people learn" (14J).*

> *"I take more opportunities to teach now. I've always found medical students refreshing because of their enthusiasm. What I consider as mundane clinical encounters to me is actually exciting and interesting to them. So I take small chunks and make it into interesting opportunities for them to learn from" (12A).*

The hypotheses supported by the two main outcomes of teaching and personal development were hypotheses 1 and 2 of hypotheses of the inquiry as shown in Table 7.14 below. This will be considered in further detail in the discussion.

**Table 7.14: Hypotheses supported by outcomes**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + | Engagement generated by the facilitator being very interactive and informative using multimodal approach | = | Increased confidence in teaching. Empowerment to utilise teaching opportunities |
| Hypothesis 2 | FD course aligned with or relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills and change in teaching approach |

## 7.5 Views on the future of FD

Aside from the realist evaluation, I was also interested in finding out from educators their views on FD in the UK because as stakeholders their views need to be harnessed to improve, refine and shape future FD development. Majority preferred short courses and workshops followed by seminar series with longitudinal courses the least popular. Having said that most educators wanted accreditation and recognition for the training they undergo hence FD developers must balance these two issues.

Educators felt that teaching should be considered a core function for most doctors and suggested that FD on teaching should be introduced early into doctors' training. However there was no consensus on the optimal time. While most felt it should be introduced immediately post-qualification, a few suggested it should be introduced during specialist training and a small number suggested it should be introduced pre-qualification i.e. during medical school training. Some voiced the opinion that a certain amount of one's ability to be a teacher may be inherent with some people having a more natural aptitude but all agreed that even this could be enhanced by

training. All agreed there needs to be a baseline standard of what it is to be a medical educator e.g. awareness of different educational techniques, and learning theories. This should be made explicit and contractual to reduce variable outcomes in teaching.

Next was the type of FD provided; educators want it tailored and targeted to teachers' needs i.e. should be developmental. Hence, they suggest there needs to be a range of different options for different people with FD tailored to the level at which they are teaching. In their view this multifaceted approach should endeavour to provide the right training for the context of the teacher. Lastly, they focused on professionalism and commented on teachers being guided and trained to have the right professional attitude and this to be developed as part of mentoring in FD.

## 7.6 Discussion

In this section, I will discuss the findings briefly but leave a detailed discussion of the CMO hypotheses till Chapter 9 when the findings from all phases had been considered. I have divided the discussion into three parts. The first part deals with the findings from the FDC interview data. In the second part, I will consider the observation findings during the ECE course and in the third part, I will discuss the CMO data in relation to the realist hypotheses exploring in further details the hypotheses that were supported or modified. I end the chapter with a summary of the important findings.

*FD coordinator data*

The key objective of the FDC role as stated in the job description was to design and deliver effective educational development opportunities for clinical teachers and develop FD activities across all clinical sites. Therein lays one of her immediate challenges as there has been no systematic needs analysis of the teachers. However, if as shown above, individual motivation is the highly predominant factor amongst educators, then FD developers have to consider tailoring and marketing FD to individual needs. This view was also echoed by educators regarding future FD programmes. In my opinion, this could be done in various ways. One is by collaborating with educators in designing the courses / interventions. Another possibility for courses like this one is to collect prior information on educators' needs, motivation and reasons for attending and then explore these during one of the sessions. This could be a 'meet the panel' type of session when all facilitators are present and participants are given the opportunity to explore these areas / needs in depth.

Another key area from the FDC data was the issue of FD evaluation. The coordinator evaluated the course mainly in two ways: a standard local module evaluation form which has a four point Likert scale (poor, satisfactory, good and excellent) and the thirty minutes module evaluation at the end of the course. The Likert scale questionnaire across the two cohorts produced information mostly at Kirkpatrick level 1 reaction (Table 2.2) which I have already discussed its many limitations in section 2.6.1. The 30 minutes module evaluation discussion at the end of the course was probably more useful as it provided qualitative data and also allowed participants to feedback issues that mattered to them. The FDC also mentioned the assessment of

participants' reflective portfolios submitted three months post course as a sort of evaluation. The argument was that since the assessments were aligned with the module learning outcomes, poor performance in an area might flag up issues to be addressed. However, in my view, at best this could only be used as a surrogate as they are a lot of variables affecting learners' assessment, which are not related to the course. Examples would be work commitment, personal / family issues and competing pressures. In my view this highlights the fact that FD evaluation to demonstrate impact and 'what works for who' is still a problem for FDC. I believe a realist evaluation is the way forward to achieve this and my study will help FDC understand how to approach the issue.

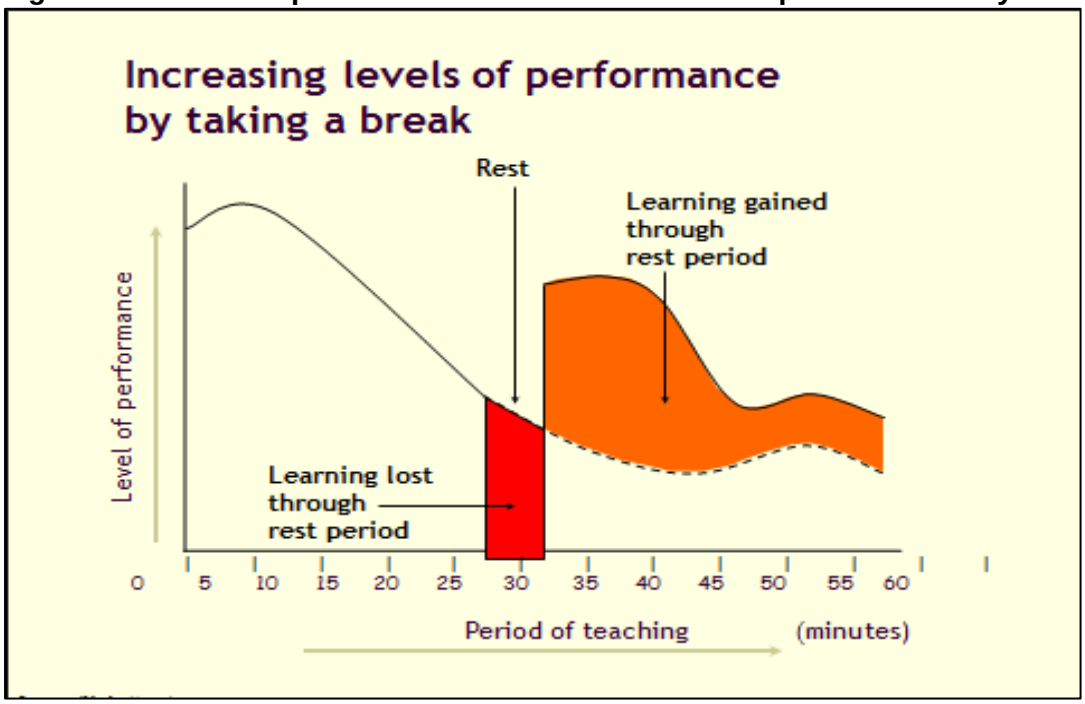
*Observation findings*

I have already discussed the rationale for adopting a non-participant observer with interaction for this phase of the study in section 4.2.2.2. Having carried out the study I identified additional advantages of this direct observation. First, I was able to integrate the observed behaviour into its context and second, I was able to gain intuitive understanding of the meaning of my data. As I have an intimate knowledge of the study and also directly experienced the FD training with the participants, I was capable of taking positions about the meaning of my data with confidence. I noted that majority of participants were engaged 31/33 (94%), scoring between three and four on most sessions without a significant pattern emerging between morning or afternoon sessions. This suggested that it was the interactivity of the sessions that mattered rather than the time of day. However, I have to admit that it was sometimes difficult during observation to decide between giving a score of 3 (partial engagement) or 2 (minimal engagement) as described in section 5.3 because the

two engagement levels are too close together. In hindsight, a 4 point scale (rather than a 5 point scale) would have bypassed the difficulty in differentiating between those two close engagement levels. On a 4 point scale, a score of 4 (full engagement) is given when all three categories descriptors (behavioural, emotional and cognitive) are observed, 3 (mostly engaged) if any two categories descriptors are observed, 2 (partial engagement) if only one category descriptor is observed and a score of 1 (no engagement) if no category descriptor is observed.

Nevertheless, a closer look at the engagement construct (Figure 7.5) showed that up to eight participants at some point felt the sessions were intense or felt unable to relax. As an observer, I noted in my contemporaneous records that up to 50% of the time participants were involved in activities and this could have been partly responsible for some of the tension or lack of relaxation experienced by some of them. I was aware that the facilitators were keen to follow Bligh's (2000) and Cantillon's (2003) advice about introducing activities for the learners after 25 minutes of teaching in order to sustain learners level of performance (Figure 7.7). While this was a good teaching strategy, most of the activities were oral / written i.e. discussion at the individual tables followed by writing on the flipchart and / or presenting discussion points to the larger group. Despite the learners being taught about learning styles including visual, auditory and kinaesthetic (VAK) right from day one, it was surprising that only the e-learning session utilised a video to break up the teaching. A video showing feedback been given or a five minute video clip demonstrating 'on the job teaching' could have easily replaced some of the writing and FDC should perhaps consider this. Potentially, this could move the engagement score from 3-4 to 4-5 and reduce the number of participants feeling tense.

**Figure 7.7: Learners performance with and without rest period of activity**




Source: Bligh (2000): What is the use of lectures? (with permission)
Graph shows how learners level of performance increases if they have a short period of learners' activity rather than continuous teaching

*CMO data and realist hypotheses*

There was good correlation between the three data sources (the observed engagement scale, the interviews during the course and the six months post course interviews) used in exploring the causal mechanisms facilitated by contexts and the outcomes. The participants who appeared disengaged to the observer also had the lowest engagement scores in their cohort and their perception of the course was that of being unproductive. On the other hand, participants who were observed as engaged also reported being engaged by the interactivity and multimodal approach of the FD. Furthermore, they had a positive perception of the course and reported good learning outcomes. Thus there was evidence supporting motivation, engagement and perception (and their respective hypotheses 1, 2, and 3) as three

important mechanisms that were key to participants learning in FD. The evidence for feedback as a mechanism was less convincing and possibly less important as feedback had a dual function of mechanism and context. I will discuss each of the mechanisms (and the hypothesis supported or modified) in the following paragraphs.

For motivation, the significant finding was that the main motivation was *individual* rather than *altruistic* with no difference between *external* and *internal* motivation. Previously, the focus had been on *internal* and *external* motivators as key drivers. Knowles et al (1998), and Merriam and Caffarella (1999), suggested that while adults are responsive to some external motivators (better jobs, promotions and higher salaries) the most potent motivators were *internal* such as the desire for increased job satisfaction, self-esteem and quality of life. However, others have argued that to construe motivation as a simple *internal* or *external* phenomenon was to deny the very complexity of the human mind (Misch, 2002; Brissette & Howes, 2010). While previous studies on motivation have varied in design, objectives, and measured outcomes, most have reported a positive correlation with motivation influencing learning, study behaviour, academic performance and success (Wilkinson et al, 2007; Sobral, 2004). However, all these studies have focused mostly on *internal* motivation and none answered the question: what is responsible for this relationship between motivation and learning and does it change over time? By analysing motivation on a bi-axial construct, I have shown the multidimensional nature of motivation. Furthermore, the combination of different types of motivation in each quadrant of the construct showed that motivation could be mutually interactive and should be viewed as a dynamic concept; thus a person could move between different types of motivation depending on the situation. As far as I am aware, this is

the first time that motivation has been explored in multidimensional constructs in relation to learning in FD. Motivation supported hypothesis 2 of the hypotheses of the inquiry.

However, while motivation was important, I found engagement activated by an interactive participatory approach to be the main causal mechanism that made the programme effective and it supports hypothesis 1. Engagement is the answer to the question posed above about the link between motivation and learning. Engagement as defined in (Box 2) is the bridge between the learner and their learning target i.e. between the inner mental states of motivational orientation and learning success. Without engagement, there is no deep learning (Hargreaves, 2006), effective teaching, meaningful outcome, real attainment or progress (Carpenter, 2010). In my opinion, the finding that engagement was the key mechanism was quite robust because I used a multidimensional observed engagement scale, which as discussed in section 4.2.2.1 has been attested to by various authors as providing a richer picture of participants' engagement. Moreover, it provided a coherent and consistent structure to the observation thereby enhancing the methodological quality. In addition it was combined with participants own description of their engagement, which was further analysed on a bi-axial construct.

Six months later, engagement was still the key mechanism responsible for learning but the contextual factor activating the engagement was now reflective practice. So in essence, I found two contextual factors that activated engagement. During the course, the participatory learning approach using a multimodal, interactive style was the context while after the course it was the reflective practice (experiential practice

with peer observation of teaching). This is probably not surprising if reflective practice is viewed as a metacognitive process in which individuals engage to explore their experiences in order to develop a greater understanding of both self and the situation (Baumeister & Vohs, 2004). According to participants, the experiential practice followed by reflection and development of action points became the driver for individual learning. This self-directed learning (*or directed self-learning*) is a type of self-regulated learning activity which has been shown to lead to a deeper approach to learning and improved performance (Baumeister & Vohs, 2004)

From the realist standpoint, identifying the contextual factors that positively influenced engagement could help FD developers incorporate them into course design and into the development of their institute's teaching culture and learning environment. Sorinola and Thistlethwaite (2013) in a systematic review reported the impact of context and environment on the success of FD initiatives in family medicine. Hence in my study it was important to identify the contextual factors activating engagement during and after the course. It was also significant that engagement was linked with participants' perception of course usefulness and relevance, as they believed the FD would improve their teaching skills. I discuss perception next.

Perception as explained in section 7.3.2.3 depends on the complex functions of the sensory nervous system, however, it is further enriched and shaped by learning, expectations and experience (Gregory, 1987; Bernstein, 2013). The philosopher Andy Clark (2011) expanded on this explaining that there can be no completely unbiased, unfiltered perception, because there is a great deal of feedback between

perception and expectation (as perceptual experiences often shape our beliefs, but those perceptions are based on existing beliefs). Therefore, it was important to understand participants' perception of the FD course as this would have an influence on the outcome postulated in hypothesis 3. All the participants except two had a positive perception of the course and reported on the usefulness, relevance, applicability of the skills gained, and the interest / enthusiasm developed following the course. This leads me on to consider outcomes in the following paragraphs.

Causal mechanisms will only lead to effective outcomes if participants are ready to change (Pawson, 2006a). So, what changes did participants notice about themselves following the FD? The personal development aspect with regards to confidence and empowerment was more dominant in participants description of the gains (and changes) from the FD compared with their description of teaching gains. While this might not be the main outcome intended according to the FDC (Box 7.1) and the webpage description of the course (section 5.3), in my view both were still related. My interpretation is that participants placed more emphasis on the confidence they needed to practice the teaching / instructional skills gained.

In evaluating FD, the *CMO* realist framework allowed me to explore, which mechanisms were most important in the learning process and which hypotheses were most apposite for the FD course. This phase of my study has shown that hypothesis 1 with engagement (M) as the key mechanism facilitated by participatory learning and reflective practice (C) leading to increased confidence in teaching and empowerment to utilise teaching opportunities (O) was the most important hypothesis. Therefore, engaging the learner is the most important thing that an

educator, teacher or facilitator should do as there is no deep learning without engagement. This is followed in my view by hypothesis 2 on motivation as this was crucial in the decision making to attend the FD and learn from it. Hypothesis 3 on positive perception was also supported while hypothesis 4 on feedback was not supported. These are summarised in Table 7.15 in the next section.

## 7.7 Summary

I now end this chapter with a brief summary of the key phase II findings. The use of multiple data sources (observations and interviews) with two similar cohorts of educators provided rich data. Inductively deriving the bi-axial constructs for the mechanisms from the participants' data gave credence to what was important to participants as adult learners. Engagement, motivation, positive perception of quality and utility were found to be important mechanisms for FD and I have explained the link between the three key mechanisms. Hypotheses 1, 2 and 3 were the relevant hypotheses supported by the phase II data (Table 7.15). Engagement seemed to be the leading causal mechanism that made the FD programme effective in the presence of the correct activating contexts (participatory approach during the course and reflective practice after the course). The key outcome was the increased confidence in teaching experienced by the educators and empowerment to utilise teaching opportunities. This phase has added another layer of explanation as it addressed FD evaluation in the immediate and short term (six months) period. It will be interesting to see if the same or different hypotheses are supported in the longer-term with more medical schools involved. This is the focus of phase III findings, which will be discussed in the next chapter.

**Table 7.15: Summary of the CMO hypotheses supported in phase II**

| | Contexts | Mechanisms | Outcomes |
|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + Engagement generated by the facilitator being very interactive and informative using multimodal approach | = Increased confidence in teaching. Empowerment to utilise teaching opportunities |
| Hypothesis 2 | FD course aligned with or relevant to the learners needs e.g. aspects of teaching they are deficient in | + Motivation to learn the necessary skills in areas of weaknesses | = Improvement in instructional skills and change in teaching approach |
| Hypothesis 3 | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + Positive perception of the value and relevance of teaching in their job and in the organisation | = Increased enthusiasm, interest and higher importance attached to teaching |

# CHAPTER 8

# PHASE III FINDINGS

## 8 Introduction

This chapter presents the findings of the interviews of the educators and the FDC at the eight medical schools as discussed under methods (Chapter 5). Phase II findings described in the previous chapter provided explanations of the hypotheses that were supported in the short term during and immediately after FD. Phase III research seeks to expose the causal links between mechanisms and outcomes as well as to understand the influence of context in a longer term period after FD. It should then be possible to see how the longer term context verified, modified, or disproved the hypotheses of the inquiry

I start the chapter with a quick review of realism, objectivity of accounts and a brief description of the FDC and educators' backgrounds to set the scene. Next, I move on to discuss the CMO findings from the interview data and the hypotheses supported, modified or disproved. Following that I discuss the views of educators on the future of FD in UK. This is followed by a brief discussion of the key findings from this phase and how the findings contributed in building the final picture of this realist study. I end the chapter with a summary of the important findings in this phase. As mentioned before, a more detailed integration and discussion of the findings from all phases is left till Chapter 9.

## 8.1 Realist accounts

Realism supports the view that entities exist independently of being perceived, or independently of our theories about them (Phillips, 1987). Maxwell (2012) further explained that while our knowledge of the world is inherently a construction from a particular perspective, there is nonetheless a real world, which can be understood in both mental and physical terms.

> ***"There is no possibility of our achieving a purely 'objective' account that is independent of particular perspectives. All knowledge is thus 'theory-laden', but this does not contradict the existence of a real world to which this knowledge refers." (Maxwell, 2012, p. vii)***

Therefore the following account has to be understood in the light of the interviewees' perspective either as FDC or educators. FDC described their experiences from the way their programmes were planned, implemented, and evaluated and the challenges they encountered. Educators described their experiences, perceptions, attitudes, changes in knowledge, skills and behaviour, barriers and facilitators to change within their own environment. Through their descriptions, I examined the mechanisms and contextual factors that may have influenced the programmes outcomes. Nevertheless, I took very seriously Robson's (2011) advice on the deficiencies of the human analyst: I was careful not to ignore information that conflicted with my ideas, and I did not discount any information.

One of the most important aspects of the realist approach is the emphasis it places on understanding the background in which interventions are introduced. Gaining evidence about background issues is not always easy as some aspects may be unintentionally missed due to the amount of data that needs to be gathered. Despite

this, I still feel it is a good approach to give useful background data on the two groups of stakeholders before moving on to describe the *CMO* findings in detail.

## 8.2 General background

All eight medical schools run a five-year undergraduate programme and three schools also run a graduate entry four-year programme for students who already have a degree. Student numbers ranged from 130 to 450 per year across the institutions. It is important to point out that even though I interviewed FDC based at the medical schools, the location of the FD programmes on teaching varied i.e. either in the medical school or centrally as I mentioned in section 6.1. All FDCs were able to provide detailed information (including content, delivery and evaluation methods) of their various FD activities which were broadly similar in terms of topic / content areas, delivery methods and evaluation pattern.

### 8.2.1 FD coordinators background and challenges

The FDC interviewed had diverse backgrounds. Only three were originally trained in a health related discipline (one physician, one dentist and one with dual qualification as a nurse and teacher) and, of the other five, two were teachers (primary / secondary school and further education college), one a scientist (biochemist), one a historian and one human resources trained. In terms of educational qualification for their current post, one had a PhD in educational technology (use of education technology to support learning and teaching), three had a master's degree in medical education, one a PGCE / a postgraduate certificate in university teaching as well as a masters in e-learning, one a postgraduate diploma in personnel management and a diploma in coaching / mentoring practice and one was doing a postgraduate

certificate in education at the time of the interview. The eighth FDC qualified as a teacher and taught in a college of further education but had no additional qualification. They had various university titles as shown in Table 8.1 and as previously discussed in section 2.1, the staff development title is more commonly used in the UK than faculty development.

**Table 8.1: Background details of the eight FD coordinators (FDC)**

| **Titles** | | |
|---|---|---|
| Director of Staff Development | | |
| Director of the Clinical Educator Programme | | |
| Director of Problem Based Learning /<br>Staff Development Officer for the Faculty of Medicine | | |
| Assistant Director of Medical Education | | |
| Clinical Education Staff Development Facilitator | | |
| Academic Lead for Staff Development | | |
| Academic Practice Advisor | | |
| Education Advisor | | |
| **Time in Post** | | **No of FDC** |
| New (less than 2 years) | (7 months; 15 months) | 2 |
| Recent (two to less than 5 years) | (2 years; 3 years) | 2 |
| Established (5 years or more) | (5, 6, 7 and 20 years) | 4 |
| **Time allocated to FD** | | |
| Full time | Full working week | 5 |
| Part time | 1 day; 2 day; 3 days | 3 |

I also found out how long the interviewees had held a position of responsibility in FD. There was an equal split between those categorised as ‘less experienced developers’ (less than five years in post) and ‘experienced developers’ (five or more years in post). This categorisation was similar to that used by Sorcinelli et al. (2006) in their survey of FDC. Furthermore, the new and recent FDC, with a range of 7 months to 3 years’ experience between them (as shown in Table 8.1), were the first in post at their respective institutions. This might reflect the recent growth of FD programmes across UK medical schools as well as the profession acknowledging

the need to make teaching skills and teaching orientation available to medical school educators.

The time allocated to FD by the coordinators varied from one day a week to a full working week (Table 8.1). The part-timers spent their time in other roles such as general practice and postgraduate medical education. There was no correlation between the medical school size in terms of the number of students' intake and the time spent on FD by the coordinators. When asked about the goal of FD, virtually all the FDC had the same goal in mind. This was best summarised by one coordinator:

**Box 8.1: Ultimate Goal of FD**

"We want the best to teach the best, we want talented high potential students to have learning experiences from talented and professional educators. The end result is excellent delivery of patient care. It's about driving the standards of teaching to drive the standards of care." FDC3

(FDC3: FDC is faculty development coordinator, number 3 is the code)

With the above goal in mind, all the FDC were keen to discuss their challenges in achieving that goal. The top five challenges are summarised in Table 8.2 below. The vastness of the job in terms of the FDC to teacher ratio (ranged from 1:220 to 1:950), and the geographical area to be covered with NHS hospitals scattered over wide areas (sometimes well over 60 miles) was mentioned by most of them. Resources both in monetary terms and personnel were another issue. Three FDC were able to state the exact budget set aside for FD (ranged from £20,000 - £40,000 / year) while the others were unable to state the budget amount or source of funding but believed it came from the service increment for teaching (SIFT) money obtained from the local education and training board (LETB – see glossary, pg. viii). Funding for FD was an important issue as it is related to the mission and governance as well as the

institutional value assigned to teaching. On average two to three people worked in the FD office (mostly not full time equivalent – FTE) and these usually included administrative staff or part-time teaching staff. Overall there exists a challenge in providing varied FD activities in the face of shrinking resources.

**Table 8.2: Challenges for FD Coordinators**

| | |
|---|---|
| Vastness / scope of the job | FDC to educator ratio, geographical spread |
| Resources | Funding, personnel. |
| Competing priorities | Time pressure, job description, outside initiatives |
| Organisational | Support, documentation of attendance, isolated role |
| Engagement | With NHS Trusts, with educators |
| Needs assessment | Gap analysis, influence on course provision and participation |

There were competing priorities and time pressure especially for those FDC who were part time with other jobs or responsibilities and with an ever increasing demand to provide more FD activities. Furthermore, external influences and recommendations, for example the recent GMC (2012) document on 'recognising and approving trainers', had increased their workload in terms of documentation and training of educators. This was aptly summarised by one FDC,

> *"I find the GMC requirement a big challenge. It says, 'We want to you to design programmes in these seven areas and the people who are going to be covered, who need to be recognised are those responsible for overseeing students' progress at each medical school.' So is that one person, or is that 20 people? It's very difficult" (FDC4).*

Engagement with hospitals and with educators, lack of organisational support, keeping track of educators (sometimes due to quick turnovers) and the feeling of isolation in the job were some of the other challenges discussed by the coordinators.

Another key challenge for the FDC was in conducting needs assessment of the faculty. While they unanimously acknowledged the importance of doing needs assessment to know where the gaps were, to help decide the type of FD provision and to encourage participation, only one FDC had formally conducted a needs assessment. The others used surrogate sources such as student feedback, educator feedback at the end of FD courses, feedback at annual medical education conferences and needs analysis reported in the literature.

### 8.2.2 Educators background

Compared with the FDC, the majority of the educators interviewed were clinicians. There were six clinicians: two in primary care (GPs) and four in secondary care (hospitals). The other two interviewees were scientists (biomedical science and neuroscience). Four of the eight had substantive university appointments and the other four had honorary appointments. Seven of the eight educators had a PGCE or a postgraduate certificate in learning and teaching in higher education and one of the seven also had a masters in medical education. The eighth educator had attended FD teacher training activities but not progressed to a formal qualification. Based on the selection criteria described in section 5.3, they were all involved in medical student teaching and had participated in FD activities within the last two years. Time in educational posts ranged from two to 14 years, while time spent teaching on a weekly basis ranged from one to four sessions (each session is a four hour block). As of the time of the interview, most educators had attended their FD programme voluntarily, even for those with university contracts. This was clearly stated by one educator,

*"Oh it's voluntary; I do feel a little… I mean I think I'd be expected to be at the curriculum workshops and conferences for example, as I have a contract with the institute of medical education and everybody else is there. You need to know what is going on… but nobody has ever told me off for not being at anything"*
*(E3 – E is educator and 3 is the code).*

## 8.3 CMO findings

In this section, I discuss the findings in relation to the contexts, mechanisms and outcomes and how the hypotheses were supported or modified.

### 8.3.1 Contexts

The contexts (Table 8.3) examined in this section were from both the educators and the FDC interviews.

**Table 8.3: Contextual factors**

| | |
|---|---|
| i. Relevance | v. Supportive educational environment |
| ii. Lack of pedagogy | vi. Access to FD |
| iii. External influences | vii. Reflective Practice |
| iv. Participatory | |

#### 8.3.1.1 Relevance

Both the FD coordinators and the educators commented on the importance of aligning the FD activities to the need of the learners. This was a strategy that was useful in improving attendance as well as stimulating further interest in teaching. For the FDC it was also the challenge of striking the right balance in designing programmes sometimes for people with different expectations, for example clinicians versus university academics. Educators affirmed that FD initiatives should help them to do their job better and should meet the requirements of their roles. Selected quotes from the two groups of stakeholders provide substantiation of the importance of relevance to educators' attendance and to FDC planning of courses.

*"I think it's got significant relevance at multiple levels really because the course dealt with large group teaching, small group teaching, one-to-one supervision and I do them all in varied contexts. The course gives you instant tips to improve from where you are. I'm terribly pragmatic, I need to get the job done therefore I need something relevant" (E6)*

*"I think it's recognising how busy educators are and try to align the courses so that they are relevant to them. And the thing that makes me more disappointed than anything is if I put a course on and it doesn't go very well because they haven't found it relevant" (FDC1)*

This context of relevance was similar to that reported by Steinert et al. (2010) as they found relevance to be one of five reasons for participation in FD. Admittedly, there is the issue of self-selection and self-assessment of needs as it has been suggested that the educators with more needs are the least likely attenders at FD. However, for those that attend, FD courses being relevant and aligned to the needs of the educator was an important contextual factor. It supports hypothesis 2 (table 8.4) and in the next section under mechanism (8.3.2), I further explore how it facilitates the motivation to attend and learn from the FD.

**Table 8.4: Hypothesis supported by relevance**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 2 | FD course aligned to the needs of the educator or relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills and change in teaching approach |

#### 8.3.1.2 Lack of pedagogy

Similar to my findings in phase II, lack of pedagogical knowledge was a contextual factor for the educators interviewed. This was a bit surprising bearing in mind that most were well established educators teaching for a number of years compared with

the educators in phase II, some of whom were quite early in their career (section 7.1). From the educators' point of view, the learning theories taught as part of the content of the course were stimulating and interesting; mostly they were glad to have some theoretical knowledge underpinning their practice. Below is a quote on lack of pedagogy.

> *"The educational theory definitely is up there as a major area after ten years of doing it by myself so to speak. I didn't really have experience in educational theories, so for me learning about them is just fascinating…. I also find it extended my knowledge so I can now say well… this is the theory that matches that" (E6)*

Pedagogy was an interesting context because when I asked the FDC which theoretical framework(s) they use in designing their programmes, only three out of the eight were able to precisely state the theory or theories that informed their programmes as constructivist and critical reflection (quote below). Two FDC affirmed that they had no specific theories that informed their programmes (quote below) and the other three were not explicit even though they claimed to think about learning theories in general.

> *"Based around the ideas of critical reflection and also getting people to understand things like social contstructivism in terms of learning … Those are sort of my touchstones" (FDC 5)*

> *"I don't suppose we thought of any theoretical frameworks at the time. It was really just, you know, it was a very pragmatic, practical thing …. We wanted to practically give people an opportunity to learn how to teach more effectively" (FDC 8)*

So in essence some FDC are 'teaching' the theories but seemed not to use theories to underpin their courses. One might argue that this could be because the FDC are so familiar with the learning theories that they tacitly run the sessions according to

these learning theories without thinking about them explicitly. This is similar to the findings from the literature review (section 2.5.1) where the majority of FD activities were not explicitly grounded in a theoretical framework (Steinert, 2012). Interestingly one of the three FDC above who was not explicit in articulating the theoretical framework used, acknowledged that the educators wanted more knowledge on learning theories as suggested in the quote below.

> *"I get that feedback and a lot of the tutors ask for more help in the theories of learning and practicality of teaching and learning …. I would have a tendency to agree that we do need to concentrate a bit more on the educational theory … in an applied sense. There seems to be a lot of feedback asking for that type of thing" (FDC 6)*

For pedagogy, the closest hypothesis is hypothesis 2, but a new hypothesis (2a) can be proposed as shown in Table 8.5. However, my view is that educators were mostly glad to have some theoretical base for the teaching they are undertaking rather than gaining a deep knowledge of learning theories as an outcome. I explored this further under outcomes in section 8.3.3.

**Table 8.5: Hypotheses supported by pedagogy**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 2 | FD course aligned to the needs of the educator or FD relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills and change in teaching approach |
| Hypothesis 2a | FD course focusing on the pedagogical knowledge and the theories of learning in which educators are deficient | + | Stimulation of interest and motivation to learn the relevant areas of weaknesses | = | Improvement in pedagogical skills or change in conception of teaching |

#### 8.3.1.3 External influences

All the FDC were strongly influenced by outside initiatives such as the GMC (2012) implementation plan on recognising and approving trainers, the HEA (2011) revised UKPSF for teaching and supporting learning in higher education, and the national policy that universities from 2013 are expected to report the numbers of staff with a PGCE or Fellow of the HEA to the Higher Education Statistics Agency (HESA). In addition, the NSS (section 6.2) highlights individual medical schools' areas of weaknesses such as student feedback. These were all part of an ongoing sector-wide professionalization of teaching and learning in UK higher education which one FDC argued was long overdue and cited instances of changes that had been made following GMC visits. GMC visits are part of the quality assurance process to review compliance of medical schools training against GMC standards. The FDC provided a particularly trenchant summary,

> *"So I looked at the HEA and the Academy of Medical Educators Professional Standards and made sure that we were covering these areas. And with this latest GMC 'Recognising and Improving Trainers' document, I'm having to go through in fine detail to check that the learning outcomes for each of the courses map to what they want us to map to. Furthermore, the GMC had said last time they visited in 2009 (which I think was part of the reason my role was created), that there needed to be more provision in terms of FD and education of teachers. Assessment and feedback was highlighted as priority; feedback being something that we had scored really badly on the NSS and so there's been moves to improve that through delivering 'Effective Feedback' workshops and setting up a 'Personal Tutor System' to make sure people are trained in how to give feedback as well as training students on how to receive and actually understand what feedback is. This latest GMC requirement is a big but necessary challenge" (FDC4)*

In contrast, only two educators considered external influence a contextual factor at the time of the interview as most had attended their FD activity in the previous two

years on a voluntary basis. There was some degree of awareness (as attested to by the comment of one educator below) that this would change, as well as recognition of the trend towards greater accountability for example with GMC revalidation. Revalidation is the process by which licensed doctors are required to demonstrate on a regular basis (five yearly in the UK) that they are up to date and fit to practise[8]. However, the overall feeling from the other educators was that this was not a significant issue in their mind or a consideration at the time of the FD. This was not totally unexpected bearing in mind the expertise of each group of stakeholder. The FDC are more aware of the changes in the external / regulatory sector as well as having direct access to NSS report for their individual schools.

> *"The FD are completely voluntary, I didn't have to do them. But because of GMC revalidation and the move towards undergraduate teachers needing to have evidence that they are capable of teaching, I saw this as an opportunity to actually do it. So it's almost future proofing myself if the university says, you can't teach students unless you have the evidence to prove that you're trained" (E4)*

This context of external influences aligned clearly with hypothesis 6 as shown in Table 8.6 and to some degree with hypothesis 7 because if attendance at FD teacher training becomes mandatory, then student feedback and evaluation of teaching may improve. Under outcomes (section 8.3.3), I will explore further the links between external influences, professionalization of teaching and career progression.

[8] GMC Revalidation: www.gmc-uk.org/doctors/revalidation

**Table 8.6: Hypotheses supported by outside initiatives and FD attendance**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 6 | Outside initiatives / external influences and demand to standardise medical educators teaching | + | Professionalization of teaching with qualification / accreditation and standards by recognised bodies or authorities | = | Career progression in teaching / improved credibility and recognition as educators |
| Hypothesis 7 | Time available to attend FD / Time available to practice what was learnt | + | Regular attendance and participation at FD and updates | = | Improvement in instructional skills; improved student feedback or student evaluation of teachers |

#### 8.3.1.4 Participatory approach

The FDC were keen to emphasise the multimodal, interactive, participatory approaches they employed in delivering their courses such as involvement of real or simulated patients, videos, learner presentations and group discussions. The intent was to appeal to various learning styles (visual, kinaesthetic and auditory), to encourage interactivity, and stimulate learning as well as provide some experiential learning as illustrated in the following quotes.

> *"We made everything as interactive as possible, so it wasn't just didactic one way teaching. We interacted, asked questions, and encouraged interactivity. For example, everybody had an opportunity to do two teaching presentations; a ten minute presentation from their own clinical area, and a non-medical presentation, something different. We videoed them, gave them individual feedback, and more recently we started sending them videos of themselves, that they could actually keep" (FDC6)*

> *"We've also used patient educators on that course which was the first time we've done that and it went really well" (FDC2)*

Similarly for the educators, they found the participatory, interactive approach very stimulating and engaging; it contributed to their learning. The participation also involved learning from and sharing experiences with peers and this increased educators' confidence in dealing with teaching scenarios as stated in the quotes below.

> *"I think one of the things that I find most helpful and engaging was the different types of teaching methods used, breaking people off into small groups and breaking things down into smaller bite size pieces and allowing a lot more to come from us the learners rather than you just been told what they think you need to know" (E4)*

> *"The course was engaging as it involved a mixture of different components. There was information given but also very good opportunity to speak to others. I think experience counts because the more experienced, the more confident you are and I think the courses helped me because it was an opportunity to talk and share experiences. I became more confident about what I would do in those situations" (E7).*

This participatory approach described by both educators and FDC supports hypothesis 1 as shown below in Table 8.7

**Table 8.7: Hypothesis supported by participatory approach**

| | Context | Mechanism | Outcome |
|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + Engagement generated by the facilitator being very interactive and informative using multimodal approach | = Increased confidence in teaching. Empowerment to utilise teaching opportunities |

#### 8.3.1.5 Supportive educational environment

Three educators discussed the context of a supportive educational environment during their attendance at FD activities. They felt that the FDC were responsive to

their needs, provided additional support when needed, and genuinely enjoyed teaching as affirmed in the quotes below. Similarly, the FDC themselves commented on the supportive environment they tried to create on their courses (sometimes successfully, sometimes with constraints) to increase learners' interest and enthusiasm.

> *"The staff development courses have been well coordinated, well organised and very supportive. You feel supported by your group and the facilitators. Because it is a supportive environment you take a little bit of that away with you and feel more enthusiastic about teaching. And you know if you had queries you could contact the coordinators and although they're very busy they would make time to support you in supporting the students" (E4)*

> *"I am pretty passionate about teaching and my experience and knowledge of teaching is that you observe it in theory, and then you observe it in practice. Then you practise it yourself whilst been observed. Then you support them all the way, which is what I have tried to do as far as staff development training is concerned. And thankfully it has worked …. the fact that I am there to back up the staff has increased their enthusiasm and interest. I think that's probably the greatest impact as far as I'm concerned" (FDC 8)*

The importance of a supportive environment to both groups of stakeholders is apparent from the above data. Even though one could argue that the educators were adult learners, well motivated to attend and learn from the FD, it was still important for them to feel the passion and support from the FDC. As I already discussed in section 2.4.2, a key factor considered to be at the heart of good teaching is passion: passion for the subject and passion for sharing the learning with learners. Passion for teaching is one of the two highest-ranking attributes reported by learners; the other was motivating and inspiring the learner (Kua et al, 2006). In my view the hypothesis mostly supported by this context is hypothesis 3 as shown in Table 8.8.

**Table 8.8: Hypotheses supported by supportive educational environment**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 3 | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Increased enthusiasm, interest and higher importance attached to teaching |

#### 8.3.1.6 Access to FD

Access to FD (in terms of location, timing, frequency, and availability) was an issue for both the educators and the FDC. Most of the medical schools have large geographical spread (3 to 4 hours driving) in terms of the affiliated hospitals the students rotate through. Therefore attending centrally located FD activities was an issue for educators. This was clearly summed up by one educator in the quote below. To mitigate against this, the FDC had taken to running FD activities at various locations, times and days (including evenings and early mornings), using web based courses and other strategies to improve access. A coordinator quote, explaining the various strategies used is also given below,

> *"The faculty is enormous because it's not just those of us who've got substantive teaching posts, it's almost every consultant [hospital specialist] … Big geographical area, 3-4 hours in the car, long distances! They have to be sure that when students are sent out to those places, they're getting consistency of teaching ... They did some surveys and I think that was the basis for putting some stuff on the web. When you've come 3 hours' drive, it's a hard day isn't it?" (E3)*

> *"The FD courses are run mainly here in the medical school, that's 70% of my time and then 30% I go out to the other hospitals and do teaching sessions there. In addition, I put things on different days ... some in the morning, some in the afternoons, different days of the week and different start times. We also have an evening programme where we have six evening lectures over the year" (FDC1)*

The hypothesis mostly supported by this context is hypothesis 3 (Table 8.9)

**Table 8.9: Hypothesis supported by access to FD**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 3 | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Increased enthusiasm, interest and higher importance attached to teaching |

#### 8.3.1.7 Reflective practice

Another important contextual factor was reflective practice. For some educators, this was through the post-course assessment (reflective portfolio, teaching practice and peer observation) and for others it was through experiential practice in their work environment. This was expressed by all eight educators as an important context in their learning two years further on from the FD they attended. Below are detailed quotes from educators,

> *"I think that with the FD programme, I've learned a lot from it mainly because you have to do quite a lot of work… think about it yourself, try things out and reflect. So it's the reflective process that's helpful, you're sort of forced to do more reflecting… personally I find that really useful as it has improved my confidence in teaching and I try out things" (E1)*

> *"I found reflective writing very hard and I remember starting to write my first portfolio and I had a one-to-one meeting with one of the facilitators. It was a short meeting to get some formative feedback on a paragraph of my portfolio…. and she said this is not reflective, I want you to think about why, what have you done, how will you change, and what will you do differently. And it was like a light bulb switched on just having that conversation, I thought, that's what I need to do. For me, that was very useful because that changed my reflective thinking and writing. I feel more confident now to support my students with their reflective writing and my teaching in general" (E7)*

The concept and merits of reflective learning have long been recognised by educationalists, and reflective learning has gained popularity as authors suggested that it provides the link between an experience and learning from that experience (Blackwell et al, 2001). Inherent in the concept of reflection is the belief that people learn by doing, or through experience (Kolb, 1984). Reflection involves mulling over, discourse with self, and exploring the experience, events, and actions using possible alternatives for explaining and hypothesising. It demonstrates an internal dialogue, arising as a result of evidence gained from personal experiences, being contextualised and informed by theory, and linking uniquely to the learner's world. An issue is viewed in several different ways (from multiple perspectives), the evidence is analysed critically and either a choice or judgment is made between actions, or, what has been discovered is integrated into a better understanding of the issue (Hatton & Smith, 1995).

In my opinion, because medical scientists and clinicians have been trained to think in particular ways, a paradigm shift may be needed for some to explore their experiences in ways that lead to new understandings hence the reason some might struggle initially with the concept. However, developing reflective practice in educators is very important as shown with the educators in phase II (section 7.4.1.5). This was also the case for this group of educators, they unanimously commented on how useful and important reflective practice was in engaging them in further learning including those that were initially uncertain of the concept. I will discuss the issue of reflective practice further in Chapter 9 but for now, hypothesis 1 is the hypothesis supported by this context (Table 8.10).

**Table 8.10: Hypothesis supported by reflective practice**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + | Engagement generated by the facilitator being very interactive and informative using multimodal approach | = | Increased confidence in teaching. Empowerment to utilise teaching opportunities |

### 8.3.2 Mechanisms

Mechanisms are the structures of a programme which lead to an effect. So what goes on within FD to influence people to change? As with any complex intervention there are various mechanisms to consider in FD. In the previous chapter, I discussed the phase II mechanisms: engagement, motivation and perception, of which engagement was the main mechanism during and six months after the course. In this section, I explore the mechanisms involved over a longer term period following an FD intervention (Table 8.11).

**Table 8.11: Mechanisms**

| |
|---|
| Engagement |
| Motivation |
| Positive Perception |
| Professionalization |
| Feedback |

#### 8.3.2.1 Engagement

As already explained in Box 7.2, engagement is the link between what learners do, the inner mental states, and learning success (Schaufeli et al, 2002). There were many examples of engagement given by educators and a pattern emerged. During the FD activity itself most educators described the interactivity, discussion with peers (different disciplines with different perspectives), shared experiences and

transferability of ideas as engaging as well as driving their learning. After the FD intervention, their engagement was fostered by reflective practice and experiential practice in their own work environment. What emerged was educators' confidence to try out various teaching strategies with their own learners and the feeling of empowerment. Two educator's examples of engagement are given below.

> *"I actually think one of the most engaging things about the programme was the interactivity and participation. There were people attending from different specialties who were teaching in different ways, different groups of students. And yet from the discussion, their stories of how they teach and the scenarios that have happened, all of us learnt from each other. This gave me increased confidence in dealing with such situations in my own teaching" (E8).*

> *"For me I need to be doing and learning it and reflecting on it. That works well for me but whether it works as well for everybody I don't know … for me to learn things I need to do, then reflect, that's how it's worked for me following the FD course. By reflecting on it, I became more confident to apply what I learnt" (E3).*

The FDC echoed the opinions of the educators in terms of engagement. They were in agreement that the interactivity during the sessions as well as the reflective practice they incorporated into the assignments (where included) were deliberately aimed at promoting learner engagement *and linked engagement to outcome*. An example is,

> *"I think it's the interactivity, people talk to one another, they feel inspired. Allowing them to reflect on their experiences in an interactive way is key... it made a big difference. People are much more engaged at different levels and I have observed a lot of them at their teaching sessions making use of interactivity and less didactic teaching. Some people are really confident and engaged to develop teaching. For me seeing them empowered and thinking about the learners they are working with is a great impact" (FDC4).*

The same FD coordinator went on to provide evidence to corroborate his/her strategy:

> *"So in his (EM) lectures, he gives learners things to prepare beforehand and then the learners talk to one another in the sessions (flipped classroom / peer instruction), they teach each other and he's facilitating the sessions more. He has done lots of research to prove that this actually makes a difference, makes the learners learn more and their assessment results improve hence I use this strategy" (FDC4).*

All the above data supports engagement as a key mechanism involved in learning with FD activities. It strongly supports hypothesis 1 (Table 8.12).

**Table 8.12: Hypothesis supported by engagement**

| | Context | Mechanism | Outcome |
|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + Engagement generated by the FD being very interactive and informative using multimodal approach | = Increased confidence in teaching. Empowerment to utilise teaching opportunities |

#### 8.3.2.2 Motivation

Motivation is generally thought to be that which gives behaviour its energy and direction (Box 7.2). It influences what people choose to do, how well, and for how long (Martin, 2007; 2008; Reeve, 2001). Therefore, with this is mind, it is understandable why motivation could be a mechanism in FD. In the phase II findings discussed in the previous chapter, motivation was found to be important for decision making about attending the FD activity as well as being important for educators to learn skills in areas where they were deficient. However, in this phase, only three educators alluded to motivation being important for their learning: one talked about being motivated by the group and the fact that he or she was learning with like-

minded people, while the other two talked about being motivated by their own need to learn as shown in the two quotes below. There was little mention of motivation by the other five educators and, while this might be due to the time expired since the FD activities attended, they gave other reasons for participating: two educators thought they participated in FD because of their university contract (though they attended voluntarily at the time), another linked it with his / her postgraduate work as an educational supervisor while one thought of it in terms of collaboration and avoiding isolation.

> *"You are in a group of people who are motivated to learn so they will help and aid you. They have their own thoughts, different backgrounds and different perspectives. Many of them are in different medical specialities but some are in nursing or other areas. They all bring different aspects, perspectives and different types of expertise hence it tends to broaden your knowledge base on the generality of education with a clinical or medical slant to it. That was very motivating and helpful for my learning" (E6).*

> *"So I was motivated, I needed for my own learning a better understanding of undergraduate teaching … rather than just doing it as I've always done as I have a real interest in medical education. So that motivation made me learn during and after the course and my teaching skills have improved a lot" (E5).*

From the FDC perspective, motivation was not a mechanism they put forward during the interviews. In fact one FDC offered a different view and shared the experience (from his / her latest FD course) on motivation to participate and learn from the course. The FDC quote below relates to some participants on the course, about one-third of the group (not the whole group).

> *"When I asked what their reason and motivation was' they said 'I'm here under sufferance, I was told I needed to do it' and then they took part enthusiastically and seemed to get something out of it. So I think yes it's valued but maybe it's not valued in anticipation*

*beforehand so much as afterwards as they later said 'oh yes I did get something out of it' but the question is: 'Is it worth it? Did they get enough out of it to have spent their precious time and I think that is a question that is individual" (FDC1)*

My interpretation of motivation is similar to that in phase II in that motivation was an important factor in decision making to attend as most educators were still attending voluntarily. However, some educators seemed a bit uncertain in describing their motivation for undergoing FD training. The underlying driver was probably still mostly individual (as the above evidence suggested) but not as clearly identified as it was in phase II. The most apposite hypothesis for motivation would still be hypothesis 2 as shown in Table 8.13 below.

**Table 8.13: Hypothesis supported by motivation**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 2 | FD course aligned to the needs of the educator or relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills and change in teaching approach |

### 8.3.2.3 Perception

Perception has been defined in Chapter 7 as an intuitive recognition or appreciation of the qualities (moral, psychological, aesthetic) associated with information from the external world. In section 7.4.2.3, I discussed how perception of FD by the educators was considered in the domains of usefulness, relevance and applicability. In this phase, educators talked about perception of the FD attended in terms of its value to their teaching. There were two distinct areas commented upon: a. individual and b. institutional. Values in this instance were defined as the importance one attributes to

oneself, other people, things, or ideas, (Saldaña, 2013). According to Saldana (1995),

> ***"The greater the personal meaning (of something to someone), the greater the personal payoff; the greater the personal payoff, the greater the personal value." (Saldana, 1995, p. 28)***

a. Value of FD to individual teaching: how the FD helped educators appreciate the importance of teaching. The majority of the educators (6 out of 8) now felt a sense of responsibility to provide excellent teaching to their students. They felt more enthusiastic and interested in teaching, and spent more time preparing and reviewing teaching session. In addition, they felt that FD was important to their own professional practice and quality improvement as they viewed it as continuous learning. Two educators even felt it was a corporate responsibility to get involved in teaching as medical students are going to be the doctors of the future.

b. Value of FD in relation to institutional perception: the educators discussed how they felt participating in FD had altered the institutional perception of their teaching. They felt they were better supported to attend other teaching courses, felt their ideas and suggestions on teaching were taken on board and that there had been acknowledgment of teaching in their promotion. However two educators also felt that, despite their participation in FD, their respective organisations could do a lot more in valuing their teaching contribution. They mentioned areas such as recognition (not necessarily monetary but in status e.g. senior tutor, module leader) and mentorship (as most had no mentor). One particular educator felt valued by the institution but at the same time felt let down

because of the lack of opportunity to discuss whether to postpone retirement and stay on teaching a year longer.

Quotes from educators illustrating their perception of the value of FD to their individual teaching are provided below followed by positive and negative comments relating to their views on the institutional perception of their teaching.

> *"I would say FD has been very important because it has really helped me to develop as a teacher…. I find having the students around is generally invigorating as they question what you're doing, make you think and improve your quality. So it's not only the learners learning, the teachers are learning too and I think if you don't teach, you don't continue to learn. The FD course has helped me to appreciate this" (E1).*

> *"I think it's almost a corporate responsibility to get involved in teaching medical students because they're going to be the doctors of the future, hence it's important to attend these teaching courses and learn from them. So it's juggling the workload with how important you feel teaching is and whether or not you attend and learn from these FD courses. Some people enjoy teaching and for me I found the FD training interesting and very useful. I have become more enthusiastic in my teaching" (E6).*

<u>Positive comments on institutional perception</u>

> *"They are very heavily into teaching and leading the teaching here is really a priority. Since doing the FD courses, they have supported me by paying for me to attend international educational meetings and when I've said 'I think we need to do this sort of teaching', they've said okay, taken it on board and gone with some of it. That credibility that I have, and being allowed to do pretty much what I wanted is good" (E1)*

> *"I think they do consider the FD training important because my contract is teaching, administration and a small amount of research … but teaching is the main one and that was a real positive because once I had completed the PGCert and became a fellow of the HEA, I really felt that they highly valued my teaching contribution when I went through the promotion process. It was a good feeling" (E7)*

Negative comments on institutional perception

> *"I don't think clinicians are recognised enough for teaching, definitely that's my opinion.I don't think it's given a focus particularly even after doing FD training. I think it's still viewed as part of the normal day-to-day practice that we try and teach the students" (E2)*

> *"I have been heavily involved with teaching and learning for many years. I have a great passion for teaching which is why I got the PGCert, did the FD training and became coordinator of SSM. I will be leaving at the end of this year but I didn't have the opportunity to say, can I stay another year…. I mean my own feeling is it would have been nice to carry on teaching if possible" (E8)*

My interpretation is that the positive perception of the value of teaching at both individual and institutional level was important to educators learning. This had to be balanced against the negative comments from the two educators but on balance there was enough support for positive perception as a mechanism. These teachers following the FD became quite enthusiastic and interested to be the best at their teaching, to continuously learn and improve, and some felt they became more credible as educators. Hypothesis 3 was the closest hypothesis supported by positive perception (Table 8.14). I will explore the issue of credibility further in the thesis under outcomes (section 8.3.3).

**Table 8.14: Hypothesis supported by perception**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 3 | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Increased enthusiasm, interest and higher importance attached to teaching |

#### 8.3.2.4 Professionalization

Professionalization was viewed along the lines of qualification and accreditation aimed at standardising medical education teaching. It was an interesting area to explore as a causal mechanism because, as discussed under context, external influences are shaping medical education and becoming increasingly important in FD consideration. Furthermore, the requirement that from 2013, UK universities are expected to report the numbers of staff with a PGCE or Fellow of HEA to HESA has focused the minds of FDC on the issue. Hence, recognition of FD courses by bodies such as the HEA and the Academy of Medical Educators (AoME) was high in the mind of most FDC, with some going as far as obtaining dual accreditation and one coordinator even suggested there should be future GMC recognition of courses. Even for courses that were not accredited by the bodies mentioned above, FDC seek accreditation from other sources such as the Royal College of Surgeons, UK for continuing medical education (CME) points.

There was no doubt that seeking accreditation for courses does involve time and effort on the part of the FDC in terms of the documentation involved. However they strongly believed it was necessary to improve teaching standards, learning as well as quality improvement. In addition, they believe it would improve attendance at FD programmes as accreditation / recognition alters educators' perception of the course as being of high quality. Lastly, they commented on the fact that newly advertised jobs for academic staff across the higher education sector are now including fellowship of the HEA or AoME as essential requirements, hence the need for their programmes to be accredited for the career prospects of the educators. Even the

FDC who had not bothered to accredit courses with HEA or AoME agreed that he or she would do so if educators are not attending because of the lack of accreditation. So even though accreditation lasts for a limited period of time and has to be renewed (three years for HEA, five years for AoME) and does involve cost to the institution (e.g. £4000 for HEA accreditation), the trend seems to be increasing support for accreditation of FD courses for the reasons mentioned above. Quotes from FDC support this view.

> *"The external agenda is what are the GMC expecting, where are the AoME and HEA setting the bar for a professional medical educator and how do we respond to that? What we've introduced for all our academics and clinical academic staff is a common baseline standard to be a teacher which has never been made explicit before. Either a PGCE or membership of a recognised professional body with clear standards on teaching (HEA or AoME). Otherwise, it will be our new medical education module programme intended to bring staff to a baseline understanding of being an educator. In addition to a certificate, we've approached both the AoME and the HEA to accredit the programme, so we will get dual recognition" (FDC3).*

> *"Universities, from 2013, are going to be expected to report how many of their staff have a PGCE or Fellow of HEA to the HESA. This national policy is pushing things and it's something we've noticed, for example, in job adverts for academic staff in the higher education sector. They are asking more and more for Fellow of HEA. Our programme is accredited with the HEA, so it enables people to gain that recognition and participants are reassured regarding standards / quality and that they would learn from our courses" (FDC7).*

> *"No we never bothered to accredit the course with the HEA. Honestly it was probably out of pure laziness, because to actually do it you have to generate so much paperwork and we thought well people are going to come on our course anyway. They are not bothered about getting accredited. What we did though is got it accredited by the Royal College of Surgeons for CME points …. If we found that people wouldn't come on our courses because it wasn't accredited with the HEA, we would make an effort to get it accredited" (FDC6).*

In contrast to FDC, only half of the educators discussed professionalization as a mechanism. They talked about professionalization in terms of qualification / accreditation, in terms of providing a sense of purpose to their learning, a sense of achievement upon completion, as well as the career recognition gained. So professionalization did act as a driver for learning as shown in the educators' quotes below. It could be argued that qualification, accreditation and standardisation is only one strand of professionalism and other aspects such as attitudes and behaviour on the job are not covered. However, bearing in mind the literature review finding of the lack of formal training for medical teachers and the call for certification of medical teachers, the qualification / accreditation aspect is important. Though not all educators compared with the FDC seemed to be aware of the external influences, this will change in time due to the national and international policies. For now there is some evidence to support professionalization as a mechanism and some support for hypothesis 6 in Table 8.15 below.

> *"Would a certificate and accreditation be important? Yes, it's recognition that you've taken part and learnt from the FD course. Knowing that you are getting something out of it provides the impetus for learning. I guess because of revalidation as well, you've got to show that you have been doing it. So yes, qualification and accreditation are very helpful to make me complete the learning and for the recognition career wise" (E2)*

> *"For me it certainly helps that you get a certificate or accreditation. You can say, 'I have been to this training, I have completed this learning event'. You have the documentary evidence when asked 'What training have you had in teaching?' It certainly did help to drive my learning and I was proud I completed it. Now I can use it to further my teaching career" (E4)*

**Table 8.15: Hypothesis supported by professionalization**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 6 | Outside initiatives / external influences and demand to standardise medical educators teaching | + | Professionalization of teaching with qualification / accreditation and standards by recognised bodies or authorities | = | Career progression in teaching / improved credibility and recognition as educators |

#### 8.3.2.5 Feedback

Feedback has been defined as information communicated to the learner that is intended to modify his or her thinking or behaviour in order to improve learning and performance (Harden & Laidlaw, 2012). As Hattie and Timperley (2007) suggested, the most powerful single thing that teachers can do to enhance achievement of their students was to provide them with feedback. In phase II, there was some evidence to support feedback as a possible mechanism for FD. However, in phase III, while most educators acknowledged feedback as being important, when asked about possible mechanisms in FD that led to changes, none of them mentioned feedback as a factor in their learning during the FD. They did not talk about feedback during the FD course or their assignments in relation to learning. One possible reason could be the interval period since the FD making it difficult for educators to recall specific instances of feedback during or after the FD in relation to their learning or it might just be the case that feedback was not a strategy utilised during the FD. The only reference to feedback came from three educators who talked about seeking feedback to improve their individual teaching (see first two quotes below). I viewed these more of an outcome since it was a change in behaviour with the educators asking for more feedback following the FD course they attended. This supports hypothesis 2 which could be rewritten as hypothesis 2b (Table 8.16).

I was expecting more comments on feedback from the FDC as I thought this would have been one of their strategies for promoting learning. However, during the FDC interviews, feedback did not feature in their explanations of the learning process as a mechanism, rather coordinators commented on providing training to educators on the importance of feedback (see third quote below).

> *"It is not knowing what did go well, it's getting that feedback. I think a lot more now about seeking feedback. So now I probably don't do any presentation without asking about feedback afterwards. I tell them, for my own development as a teacher I want the feedback" (E5).*

> *"One thing that I find quite frustrating is I have tried to get feedback through the faculty on the teaching that I do because I know that the students do feedback. I've asked, 'Can I have some feedback on my component' and the response is 'Oh there are no issues'…. without knowing how your audience perceives the teaching, whether it's useful or not, you can't improve it" (E4).*

> *"Two years ago we did very badly in the NSS so of course there was a big drive to improve feedback to students, and so we did an awful lot on feedback during the FD courses in the following year" (FDC1).*

My interpretation of the above is that while educators understood the importance of feedback, and some went out of their way to seek feedback following the FD (i.e. as an outcome – hypothesis 2b), they did not consider it a mechanism that underpinned their learning during the FD. In my view, hypothesis 4 (Table 8.16) did not seem to be an important one with this group of educators.

**Table 8.16: Feedback supports hypothesis 2b but not hypothesis 4**

| | Context | Mechanism | Outcome |
|---|---|---|---|
| Hypothesis 4 | FD using an iterative cycle of training, changes to course design and continuous dialogue with stakeholders | + Feedback to the educators during the FD and / or assignments | = Improved teaching performance |
| Hypothesis 2b | FD course providing training on feedback skill relevant to the need of the educators | + Motivation to learn the necessary skill in areas of weaknesses | = Improvement in feedback skills / change in educators behaviour e.g. seeking feedback |

### 8.3.3 Outcomes

When asked about the outcomes of the FD course they attended, educators mentioned various outcomes. I have labelled the top four outcomes as the 4Cs: confidence, competence, credibility and career progression. The outcomes can also be broadly grouped into two domains: teaching development and personal development as shown in Table 8.17.

**Table 8.17: Faculty Development Outcomes – The Four Cs**

| Outcomes | Descriptors | Domains |
|---|---|---|
| **Confidence** | Confidence as a teacher; empowerment to utilise teaching opportunities | *Personal Development* |
| **Competence** | Improvement in instructional skills; change in teaching approach | *Teacher Development* |
| **Credibility** | Perceived credibility as an educator (by students and peers); higher importance attached to teaching | *Teacher Development* |
| **Career progression** | Teaching career progression; becoming FD tutors, further educational qualifications | *Personal Development* |

#### 8.3.3.1 Confidence in teaching

Confidence has already been defined in Chapter 7 as the ability to be certain. Similar to the findings in phase II, it was in this domain of personal development that educators noticed the greatest impact. They reported improved confidence in their teaching which wasn't content related. They also felt empowered to recognise and utilise teaching opportunities and a definite shift in attitude. They voluntarily offer more teaching opportunities to learners, made themselves more available to teach and encouraged learners to seek teaching opportunities. Educators described the knowledge they gained from the FD and the confidence and performance that followed from that. They also talked about confidence in terms of discussion with colleagues on teaching, been able to voice opinions on teaching issues, handling various teaching scenarios and difficult classroom dynamics. What was significant about this was that all eight educators mentioned improved confidence in teaching with over 15 references made. Some examples are given below.

> *"It's definitely made me more confident. I am certainly more comfortable doing what I do now than before because I make my teaching more interesting by using different techniques. And because I feel more comfortable, more confident, more relaxed, more natural, that has a positive impact on the learners as well. It has a positive impact on me because I don't have hypertensive episodes before teaching…. both learners and myself enjoy it more which is all positive outcome" (E4).*

> *"I think knowledge itself potentially gives both competence and confidence, as once you understand an area better, or have been exposed to an area more, you start to appreciate more of the information …. you develop competences and feel more confident in yourself and the way you go about something. In a sense it is moving towards a more ideal performance stage where you can be more productive, more contributory and confidence is helpful. Some people have more natural confidence than others; I gained my confidence following the FD through feeling that I am beginning to do*

*a job well and consistent with professional views / standards. Also I am able to have more meaningful discussions with colleagues about education…. I feel at ease and can contribute constructively" (E6)*

*"I am more confident now and I think the courses helped me greatly. My perception was that it speeded up the process of me becoming more confident in how to deal with certain groups of students in certain circumstances. For example, if students are not talking, I feel a lot more confident to try and get them to engage now, whereas years ago I would probably just carry on talking because I wouldn't know what else to do. But now I can handle that sort of classroom dynamics. Overall I feel more confident in my teaching" (E7).*

Confidence in teaching was the single most important outcome reported by the educators. This is similar to the outcome reported in phase II (section 7.4.3) and it supports hypothesis 1 (Table 8.18).

**Table 8.18: Hypothesis supported by confidence**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + | Engagement generated by the facilitator being very interactive and informative using multimodal approach | = | Increased confidence in teaching. Empowerment to utilise teaching opportunities |

#### 8.3.3.2 Competence

Competence is defined as "a combination of attributes underlying aspects of successful professional performance" (Gonczi et al, 1999, p.182). It is concerned with application of professional knowledge and skills within the workplace and is underpinned by teachers' professional values. Competency is also viewed as the ability to apply knowledge and skills to produce a required outcome i.e. it is performance based (Trinder, 2008). It places a high premium on performance of tasks and activities as well as emphasise behavioural measures that depend on integrating knowledge and skills derived from an aggregate of educational

experiences. The educators in phase III described competence in terms of changes in teaching approach (e.g. more time planning teaching, understanding the curriculum, creating a learning environment), changes in attitudes and behaviour (e.g. being more open minded, not shying away from teaching), and application of the knowledge and skills gained from the FD course. Quotes from educators reflecting these aspects are provided below.

> *"I think doing the FD course is about competence. Now I am more open to different ways of teaching, I have a clearer view on how much information to put in, what the different modalities of putting information are, so yes, the FD has made a difference. It is the multitude of small hints and tips all of which aggregate together to give a changed approach but I also find it extended my knowledge and it challenged me to extend my teaching approach. For example integrating questioning, activating prior knowledge, techniques like snowballing and pyramids, I had not really used those before the FD but now I do. I think I have got a broader and more sophisticated template on which to work and I am more open minded" (E6).*

> *"If anything, I used to shy away from teaching and if I did, I would just go along and teach anyhow, there was no structure to it. But since attending the FD course, I have developed a structure. I want to know the curriculum and what they should know then I want to know what they want to learn. And then, it's teaching that in a much more interactive way…. toing and froing of information from one person to the other. I'm better at that sort of thing; I get better feedback that way. So I have really changed" (E2).*

This outcome as described by the educators aligns with hypothesis 2 as shown in Table 8.19. The hypothesis can be modified slightly as hypothesis 2c to reflect all components of the outcome.

**Table 8.19: Hypothesis supported by competence**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 2 | FD course aligned to the needs of the educator or FD relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills |
| Hypothesis 2c | FD course aligned to the needs of the educator or FD relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Competence as shown by improvement in instructional skills and change in teaching approach |

#### 8.3.3.3 Credibility

Credibility here is focused on the credibility of the teacher. Credibility has been defined as the quality, capability, or power to elicit or inspire belief. Many authors have characterised credibility as having three domains: expertise, trustworthiness and dynamism (Haskins, 2000; McCrosky & Young, 1981). Expertise refers to mastery of knowledge matter. Dynamism refers to how a teacher presents his or herself in interactions with learners. To be dynamic, teachers must present material in a manner that excites and inspires the students. Diversity in delivery of instruction, use of technology, games and discussions are all methods that lend diversity to presentation of material. Trustworthiness refers to whether or not the learners perceive the teacher has their best interest / success at heart (Haskins, 2000).

The educators believed that participating in FD led to more credibility with their students and peers. Credibility as alluded to above has both objective as well subjective domains. While expertise is mostly objective, dynamism can be both objective and subjective and trustworthiness is mostly subjective. This makes

credibility quite a difficult measure, but despite this, educators were able to articulate why they thought they became more credible after the FD and their view of the importance of credibility in their teaching. They explained that credibility was a step following on from the confidence they've gained. They described credibility in terms of demonstrating teaching knowledge and expertise, relationship with other colleagues, perception of how they are viewed by the educational establishment *and* ability to influence discussions / decisions on teaching. Below are some quotes on credibility.

> *"Doing the FD for me was about two things; it was about confidence and credibility. In terms of confidence it is been more assured of the things that I say; I can say this is a way that I found practically helpful but is also consistent with this theory and framework. This I think gives me more credibility in teaching and also credibility in how others value my teaching" (E7).*

> *"In a sense by doing the PGCE I am sort of seen as making efforts to make myself more one of the educational establishment. And so in that way has enhanced my reputation within the deanery or LETB. I am seen as having more credibility by engaging with that process…. something that those people value… it is helpful in my linkages, it has put me in a stronger position to negotiate with more senior educational colleagues. I find the interactions with others work better and you can make things happen that wouldn't otherwise happen…. so it's just positive" (E6).*

> *"The main usefulness for me of the staff development is it gives the credibility and confidence that you're not untrained, inexperienced and you actually are doing the right thing. It is important if you really value the teaching you are doing." (E1).*

In my view, the most relevant hypothesis is hypothesis 3 which can be slightly modified to become hypothesis 3b as shown in Table 8.20.

**Table 8.20: Hypotheses supported by credibility**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 3 | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Increased enthusiasm, interest and higher importance attached to teaching. |
| Hypothesis 3b | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Credibility as a teacher / increased enthusiasm, interest and higher importance attached to teaching. |

#### 8.3.3.4 Career progression

In career progression, educators described various opportunities that became available following the FD programmes such as being promoted, becoming module leaders and facilitators on teaching courses. Others demonstrated the scholarship of teaching as described by Boyer (1990) by making their teaching available to others to use and critically appraise, develop academic writing, publications and production of teaching guides. Educators were very quick to point out that without the FD, they would not have taken up some of these roles. Moreover, since taken on these roles, they seemed to be genuinely stimulated, become more passionate and enjoy the opportunities provided by these roles. Some quotes reflecting these are given below.

> *"Doing the PGCE has positively changed my career at two levels. First, it enabled me to teach on the other PGCE programmes. I was invited on the programme committee for CLAD (Centre for Learning and Academic Development). I was then asked to be part of the PGCert in teaching and learning across the university. Last year, I became one of the facilitators on the PCAP (postgraduate certificate in academic practice) and that was great. I love doing those teachings because I get to teach and meet other academics across the university, discuss teaching, support them and do their assessments. Second, teaching on the PGCE / PCAP really supported my promotion; I was promoted last year to senior lecturer.*

*I am on a two-legged contract (teaching / administration). For my teaching element, the FD really supported my promotion. I couldn't have done all these without my PGCE" (E7).*

*"I'll show you this module guide. I was approached last year to be the lead for this module. If I had not done the PGCME I would have turned it down but because I felt I have sufficient knowledge and understanding I took it on ... Furthermore, the PGCME has improved my academic writing including VLE (virtual learning environment) writing as people have commented 'you know what you write about' …. So far I have had three publications accepted out of four. I also produced a good practice guide, which with the peer observation template is used to critically appraise learning sessions. The PGCME has made me more critical, because when I see teaching not being done appropriately, I suggest it could be done better" (E5).*

*"Doing the FD has certainly opened doors. I am now allowed to be a clinical educational supervisor as part of my job and I'm leading prescribing teaching at the university which is great. I have become a facilitator on one of the FD courses and that's made me more interested, I enjoy it. Showing that you have done the FD training is attractive to an employer as they're looking at what you can do" (E2).*

There were also comments from the FDC describing teaching career progression as an outcome as they saw former learners coming back to teach on their FD programmes as affirmed in the quote below.

*"I mean another obvious way that I see things is when you're training people to be teachers and then two or three years later they start teaching on your programme, I find that very fulfilling and so far we have had three people come back to teach on our courses" (FDC1).*

The hypothesis supported by career progression is hypothesis 6 (Table 8.21).

**Table 8.21: Hypothesis supported by career progression**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 6 | Outside initiatives / external influences and demand to standardise medical educators teaching | + | Professionalization of teaching with qualification / accreditation and standards by recognised bodies or authorities | = | Career progression in teaching / improved recognition as educators |

## 8.4 Views on the future of FD

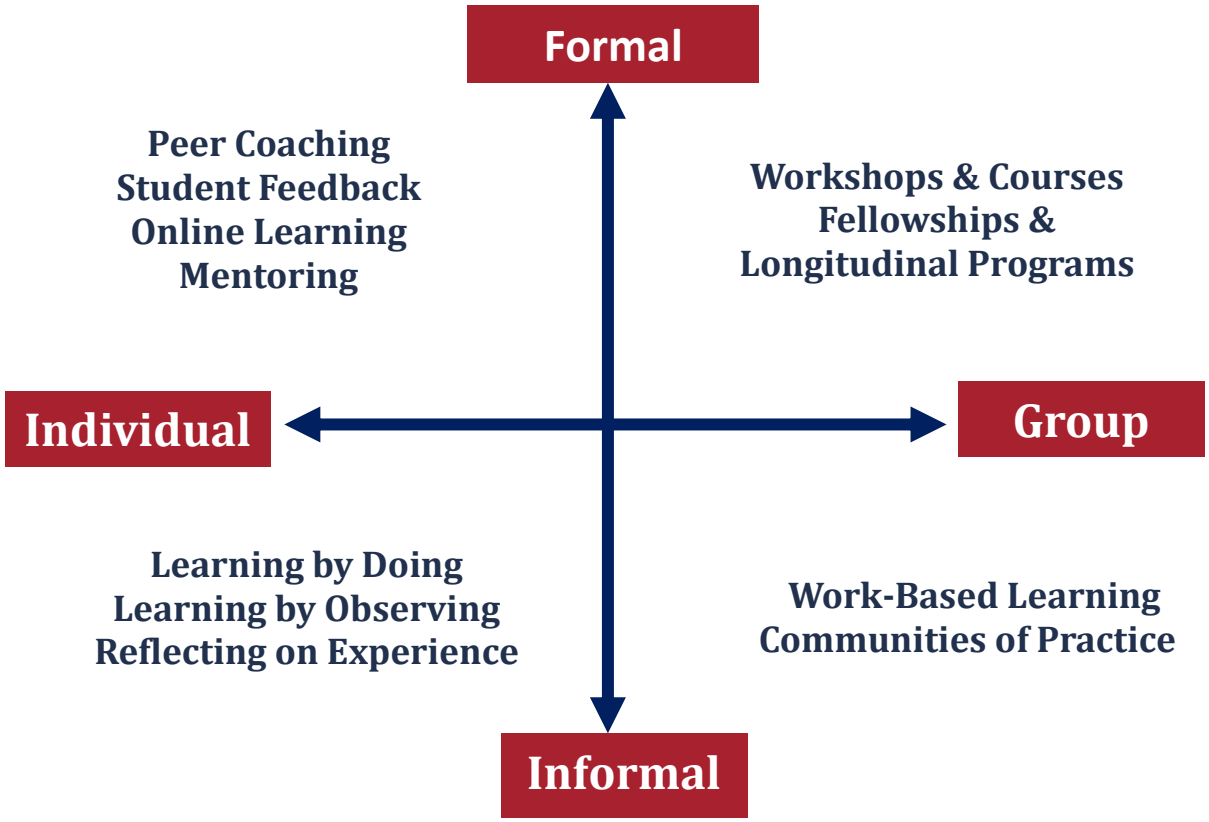
Apart from the CMO findings in relation to realist evaluation, I was also keen to find out from the two sets of stakeholders, their views on the future of FD in the UK as this was still unexplored. I wanted to know where they saw FD in the next five years or so especially as both groups had already commented on the challenges facing FD. In the following paragraphs, I have summarised the comments from both educators and FDCs.

All stakeholders suggested that a baseline standard for a medical educator need to be made explicit and contractual. The contractual component will remove the current issues of voluntary participation as well as provide clarity within the muddy pool of two groups; the academic staff with substantive contracts and the clinical staff with honorary contracts. To balance that, both groups of stakeholders agreed that the types of FD provided have to be re-examined. In Chapter 2, I discussed the various types of FD and my phase III findings have confirmed what was reported in the literature. The formal-group quadrant is the FD most commonly provided, while the informal-group quadrant was less commonly utilised (Figure 8.1). Of the other quadrants, mentoring was patchy and distinctly lacking in most places. Similarly, as reported in the literature and in my phase II findings (section 7.5) most educators in this study preferred workshops or short courses because of their inherent flexibility.

However, the FDC suggested that FD emphasis needs to move away from being event led: it is about a process, and some of that process is around building a culture of learning which is learner-centred. They envisage more local ownership of FD using a distributive model based on needs assessment and local context so that

educators have more ownership of how their FD is shaped and delivered. The educators agreed but went further to clarify that FD on teaching needs to be targeted and specific to educators' needs and be context related. They want to see FD training valued more at institutional level with incorporation into the curriculum for those in training as doctors and contractual for those established (as mentioned above).

**Figure 8.1: Types of Faculty Development Activities**




(Adapted with permission from Yvonne Steinert's presentation at the 1st International Conference on Faculty Development in the Health Professions, Canada, 2011)

On the issue of maintaining teaching standards, the FDC suggested introducing a robust annual educational appraisal for all clinical educators that teach medical

students on the basis that annual appraisal already exists for educators with substantive university posts. This should be supported by a reflective teaching portfolio which is evidence-based. However, they accept that this would require organisational change, resources, as well as a cultural change especially for the clinical educators who already undergo annual appraisals on the clinical component of their work. They stressed that the key is not to be too imposing, as educators should feel this is a fundamental part of their professional development in teaching. On their part, educators suggested that FD should become more of an iterative process with refreshers every 12 months to consolidate learning, prevent stagnation and keep momentum. The aim here is to promote the idea of continuing not just being lifelong learners but lifelong teachers. Both groups acknowledged that for this yearly process to happen, FD needs to be well-resourced. I have summarised the key points and views on FD future in Box 8.2.

| **Box 8.2: Views on Future of FD** |
|---|
| 1. Need for a baseline standard for medical educators to be made explicit and contractual |
| 2. Future FD opportunities have to be tailored and targeted to educators needs as well as the provision of advanced programs (contextual or individualised) |
| 3. FD needs to be valued more at the institutional level and incorporated into the curriculum for those in training |
| 4. To maintain teaching standards, there should be annual educational appraisals for educators (clinical / academic). This will require resources (funding / time) |
| 5. Annual FD refreshers for educators to consolidate learning and promote the concept of life-long teachers and learners |

## 8.5 Discussion

The discussion here will be limited to a consideration of the phase III data in this chapter. The phase III aim was to support, modify or invalidate the hypotheses of the inquiry and much useful stakeholder information was collected to this end. This

section will focus on my examination and interpretation of the phase III findings. A discussion of all the data collected in the study to answer the research question will be considered in Chapter 9.

A basic feature of realist research (section 3.1.5) is to seek substantial connections or patterns among phenomena rather than formal associations or regularities i.e. the links between mechanisms, outcomes and contexts (Sayer, 2000). Each set of stakeholders brought their individual expertise to the interviews and were able to comment on different aspects of the various hypotheses. However, data were not available for every hypothesis. For example, there were no data for hypothesis 5 on collaboration (Table 3.2).

Furthermore, the weighting of the data had to be considered within the bounds of realist research in that different stakeholders bring with them different knowledge and expertise (Pawson & Tilley, 1997) which affects the balance of data under the hypotheses. For example, there were far more data provided by both groups of stakeholders on the perception of the value of teaching either positively or negatively and at individual or institutional level because all respondents had an interest / view on how FD was perceived (hypothesis 3, section 8.3.2.3). In contrast, it was mostly FDC that spoke extensively about external pressures and outside initiatives (hypothesis 6, section 8.3.2.4). It is, therefore, not the volume of data only that gives weight to conclusions about the hypotheses; it is just as important to acknowledge that data were provided by respondents who were knowledgeable in those areas.

One key finding of this phase is that even though the medical schools were purposively sampled as reported in Chapter 6, the goals, challenges, contents and delivery of FD were quite similar as described by the FDCs and the educators. Even schools classified as having poor FD webpages were found to have detailed FD activities similar to schools with excellent FD webpages. This again highlights the limitation of webpages and the issue of webmasters maintain the pages as discussed in section 6.4. In the following paragraphs, I will discuss the *CMO* findings followed by the four hypotheses supported.

*Contexts*

The contextual factors of participatory approach, reflective practice and lack of pedagogy were quite similar to those reported in phase II findings in Chapter 7 which gave further credence to these factors as being important in FD learning. The other key contexts were supportive educational environment and access to FD, which facilitated positive perception of FD. Positive perception (as discussed below under mechanisms) took on a higher significance in this phase hence it was vital to understand its facilitating contexts. Relevance was important to motivation as both educators and FDC commented on the need for relevance of the FD to the teaching the educators are delivering. Both set of stakeholders understood its importance in motivating learning.

*Mechanisms*

Engagement, motivation and positive perception were again found to be important just as reported in phase II findings (Chapter 7). This again strengthened the argument for these three mechanisms as being important in FD. However, similar to

phase II, while motivation was important in the decision making to attend the FD and learn from it (majority attended voluntarily), it was engagement (facilitated by reflective practice), which brought about most of the learning. Interesting though was the fact that positive perception as a mechanism took on a higher significance during this longer term follow-up. The positive perception of the value of FD for educators' teaching as well as its contribution to the institutional perception of their role was highly commented upon by both groups of stakeholders but more so by the educators as they made an important connection to credibility as an outcome.

So in effect, in phase II immediate and short term (six months) period, the two key mechanisms were engagement and motivation in that order. However, in phase III with a longer term follow-up, the two mechanisms were engagement and positive perception in that order. A possible explanation for the higher significance of perception might be that educators have had a longer time to reflect and experience the value of FD at both individual and institutional level compared with phase II. Perhaps the best summary of phase III mechanisms was provided by one educator,

> ***"I think it's been the engagement in terms of feeling that we actually belong and we are part of the process rather than just stepping in and delivering something and then stepping out again. I think we understand where our teaching fits and in particular we can see our students developing and meeting those learning outcomes" (E7).***

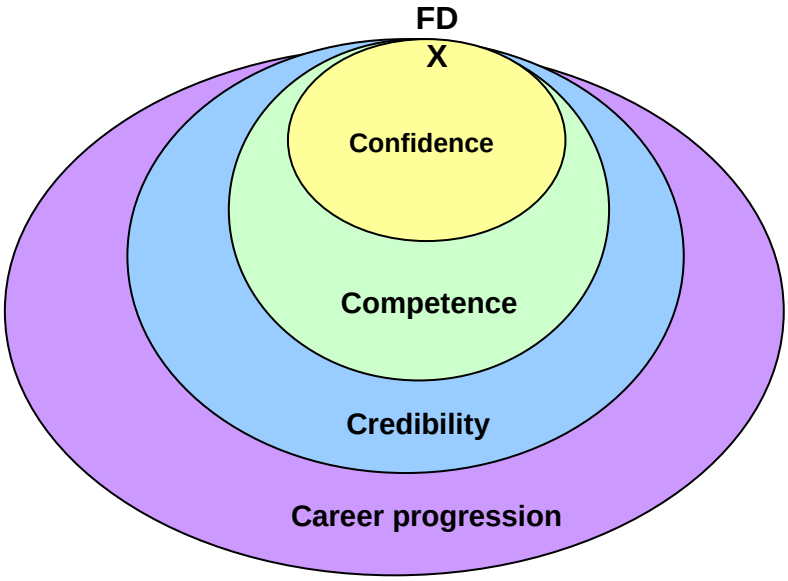
Feedback was also explored as a mechanism but the evidence for this was very limited. As in phase II, feedback was not strongly supported as a mechanism. Feedback in phase III was more of an outcome as a behavioural change with educators seeking more feedback following the FD. The issue of feedback as a mechanism or an outcome is discussed further in Chapter 9. Professionalization, the

fourth key mechanism identified in this phase was important because of the ongoing sector-wide national and international call for professionalization of medical education teaching. This will be discussed in more detail in the next chapter.

*Outcomes*

Phase III provided data on hypotheses 1, 2, 3 and 6, which as I have shown in Tables 8.19 and 8.20 could be modified slightly to take into account the reported outcomes. My conceptual view of the outcomes is represented diagrammatically as a ripple effect in Figure 8.2 with FD at the focal 12 o' clock position, and an expanding ring of outcomes emanating from it. The four main outcomes deserve further consideration; hence I have discussed them in some detail below.

**Figure 8.2: Faculty Development Outcomes – The Four Cs**



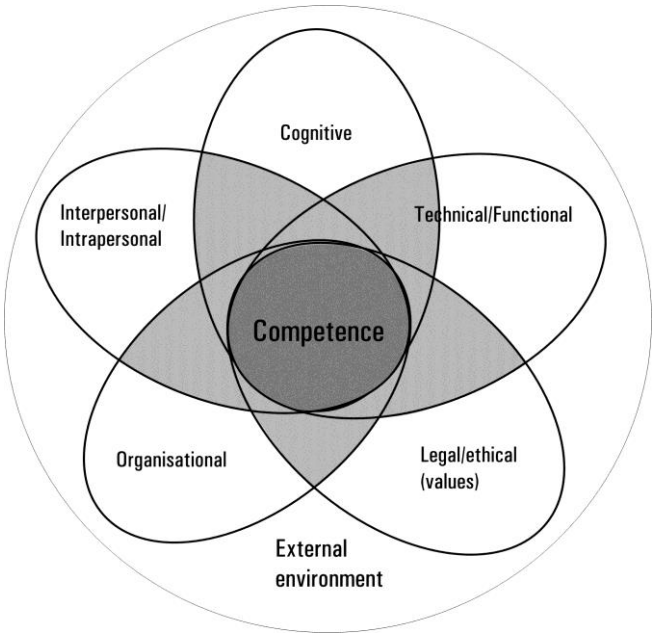

The innermost ring or core outcome was confidence, which as shown in phase II findings was the most important outcome in the short term. In phase III, it was also

confirmed as being the most important outcome in the longer term period. Educators described confidence in teaching in terms of using various teaching methods, utilising teaching opportunities, handling difficult classroom scenarios and discussion with peers. However, it is also important to highlight the fact that it was not only the educators that recognised the confidence in teaching, this was also observed by the FDC as one FDC noted,

> **"I have observed a lot of them at their teaching sessions making use of interactivity and less didactic teaching. I think we've definitely got more people engaged at different levels and some people are really confident and engaged to develop teaching. For me seeing them empowered…" (FDC4)**

The second outcome ring was competence, which could be viewed as a performance-based application of professional knowledge, skills and values to produce a required outcome (Trinder, 2008). This performance-based model emphasises the behavioural aspects of the teacher when integrating educational knowledge and skills, and teaching competency is expected to develop from the three components: education, training and experience. However, other authors have conceptualised an integrated, multi-levelled model of competency to provide a more holistic description (Cheetham & Chivers, 1996; 1998; Torr, 2008). Torr (2008) described five domains of competence: *cognitive* (knowledge for professional practice); *technical* (psychomotor skills); *legal and ethical* (values and knowledge of professional, social and cultural norms); *organisational* (ability to manage self and others); *intra and interpersonal* (communication skills, interpersonal skills, self-awareness and team work) as shown in Figure 8.3.

**Figure 8.3: The five domains of competence**




(Reproduced with permission from Amanda Torr, 2008)

Whichever competency model is favoured, for educators, becoming competent by having the requisite knowledge and skills from attending the FD and applying this with the right attitude and values in their own teaching context was important. There was evidence to affirm that educators became quite influenced by their experiences on the FD programme and the reflective practice afterwards and became more self-actualised with a greater sense of efficacy. This led them to take action and persist in their effort required to bring about successful implementation of change. My data showed that the change in teaching approaches for these educators have been multidimensional: the design of new instructional strategies, the use of new teaching approaches as well as alteration of beliefs (pedagogical assumptions / learning theories) about teaching and learning.

The third outcome reported by the educators was credibility, which was built on the back of the two preceding outcomes. Credibility is a concept that had been scrutinised since the time of Aristotle. Aristotle referred to credibility as *'ethos'* and suggested that it consisted of three dimensions: intelligence, character, and good will Kennedy (1991). He believed that those three dimensions of credibility were perceptual sources of influence on a receiver. Grounded in the Aristotle's concept of ethos, other authors have described credibility in various ways such as 'believability' (McCrosky & Young, 1981) and being situated via the 'eye of the beholder' (Kough, 1997). Robinson and King (2002) suggested that the credibility of an individual is primarily dependent upon how they are perceived by others and later added that credibility seemed most necessary to attain within the classroom setting (Robinson & King, 2002).

The educators in this study believed that following the FD and their reflective practice, their knowledge and expertise in teaching was better and together with the behavioural change in teaching approach, they improved their credibility with their learners. They expressed a genuine honesty of intent to understand and help their students more and all this played into their teacher credibility. The literature provided support for improving credibility in this manner. William Haskins (2000), a professor of communication at McKendree University in Illinois, in a seminal paper on the subject of teacher credibility wrote,

> ***"Whether at the conscious or unconscious level, a student's perception of the teacher's ethos, or speaker's character, has an important impact on how he or she will react to the teacher and how effective the teacher will be in the classroom." (Haskins, 2000, p. 1)***

He had no doubt of the vital role of credibility in educational success and said that perception of credibility is a critical factor in how learners and teachers connect with each other. He described three characteristics of teacher credibility as trustworthiness, expertise and dynamism. Trustworthiness was achieved by the pedagogical communication process that teachers employed with their learners; the learners must feel the teacher has their best interests at heart and truly cares about their success in class. Teachers who are consistent and fair, include all learners in activities and treat all equally are more likely to develop trustworthiness (Haskins, 2000). Expertise was the second key component of credibility. A teacher must not only have mastered their particular subject area, but must also deliver it in a meaningful way using good classroom management skills, the ability to answer questions and the capacity to explain complex material in a way learners could understand. The third component of credibility was dynamism. Teachers must be dynamic in how they interact with their class and present their material in an exciting and engaging way, using a diverse range of techniques. If a teacher lacks charisma or is unenthusiastic about the topic, they will lose credibility with learners. According to Haskins (2000),

> ***"Teachers need to appear not only in control of their environment, but energised by it. To look bored or distant invites communication disaster. Such impressions can quickly evaporate any feelings of excitement students have for their teachers or class content." (Haskins, 2000, p. 5)***

Credibility was therefore a very important outcome for the educators as they felt that FD gave them the platform to build their credibility with their learners, peers and the

educational establishment with *the* ability to influence discussions / decisions on teaching.

The last outcome ring in Fig 8.2 is teaching career progression which understandably was very important to educators especially earlier on in their careers. They talked about FD opening career doors, promotions and leading to various career opportunities. One could also consider career progression a culmination of the preceding three outcomes i.e. confidence, competence, and credibility all contributing to the educators career progression.

One could argue that majority of the outcomes were based on self-assessment with some supported by peer information from the FDC (see under confidence above). Self-assessment was defined by Colthart et al. (2008, p. 141) as "a personal evaluation of one's professional attributes and abilities against perceived norms". The reliability and weaknesses associated with self and peer assessments are well documented (Eva & Regehr, 2005; 2011). Accurate self-assessment may be impaired by the tendency of a person to wish to present themselves in a good light and to mask deficiencies, the social desirability response (Hartman & Nelson, 1992) and sometimes there may be a tendency for peers to be more lenient towards people they like (Falchikov, 1995). However, while it is often appropriate to consider sources other than self-assessment in an evaluation design, for FD purposes it is particularly important for faculty to make their own judgements, since adults will reject programmes they view as irrelevant (Bland & Froberg, 1982). Furthermore, Colthart and colleagues (2008) in a systematic review concluded that the accuracy of self-assessment as a measure of performance may in fact be no worse (and no

better) than any other single judgement of competence. In the final analysis, according to Bennett et al. (1992, p. 117), it is the actions of the individual that count. So in my view, it was appropriate that the outcomes reported were the ones that were relevant and uppermost in the mind of the educators in relation to their learning from FD. The difficulty with evaluating teaching was aptly summed up by Stronge et al. (2011),

> ***"Teacher quality is a complex phenomenon, and there is little consensus on what it is or how to measure it. In fact, there is considerable debate as to whether we should judge teachers based on teacher inputs (e.g., qualifications), the teaching process (e.g., instructional practices), the product of teaching (e.g., effects on student learning), or a composite of these elements" (Stronge et al, 2011, p. 340)***

*Hypotheses supported*

Having considered the outcomes, the next consideration was the hypotheses that were supported or modified by the CMO findings. Four hypotheses were strongly supported by the findings: hypotheses 1, 2c, 3b, and 6 as shown below in Table 8.22. I will suggest that for both FDC and educators, hypothesis 1 whereby engagement was facilitated by reflective practice leading to increased confidence in teaching may be the most important as this was the one that immediately followed the FD; it is the core that became a ripple source for the other outcomes.

**Table 8.22: Four hypotheses supported by the CMO findings**

| | Contexts | | Mechanisms | | Outcomes |
|---|---|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + | Engagement generated by the facilitator being very interactive and informative using multimodal approach | = | Increased confidence in teaching. Empowerment to utilise teaching opportunities |
| Hypothesis 2c | FD course aligned to the needs of the educator or FD relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Competence as shown by improvement in instructional skills and change in teaching approach |
| Hypothesis 3b | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Credibility as a teacher / increased enthusiasm, interest and higher importance attached to teaching |
| Hypothesis 6 | Outside initiatives / external influences and demand to standardise medical educators teaching | + | Professionalization of teaching with qualification/ accreditation and standards by recognised bodies or authorities | = | Career progression in teaching / improved recognition as educators |

## 8.6 Summary

This chapter formed the core of the research findings with interviews of both sets of stakeholders. The FDC interviewed had diverse backgrounds with an even split of those with less and those with more than five years experience of providing FD but all had similar challenges and the same goal for FD. The educators were mostly clinicians with an even split between substantive and honorary university appointments. While the use of a realist approach was a challenging task, the framework did provide a tool with which to evaluate FD activities. Realism assumes that each time a mechanism operates it can meet with success or failure depending on the context in which it operates. The four mechanisms in action in FD

(engagement, motivation, positive perception, professionalization) and their facilitating contexts are very important for FD developers to understand. Also highly significant were the multidimensional outcomes reported by the educators (confidence, competence, credibility and career progression), which go a long way towards justifying the need of FD for teachers. Perhaps this was best summarised by the following statement,

> ***"The best of teachers, and there are many of them, know the importance of teaching. They not only teach to the mind, but teach to the heart as well." Timothy Sullivan - President/College of William and Mary, 1992-2005. (Hall & Walck, 2007, p. 1).***

In the next chapter, I will discuss the findings from all the three phases and the cumulative understanding of the CMO hypotheses provided by all the findings considered together.

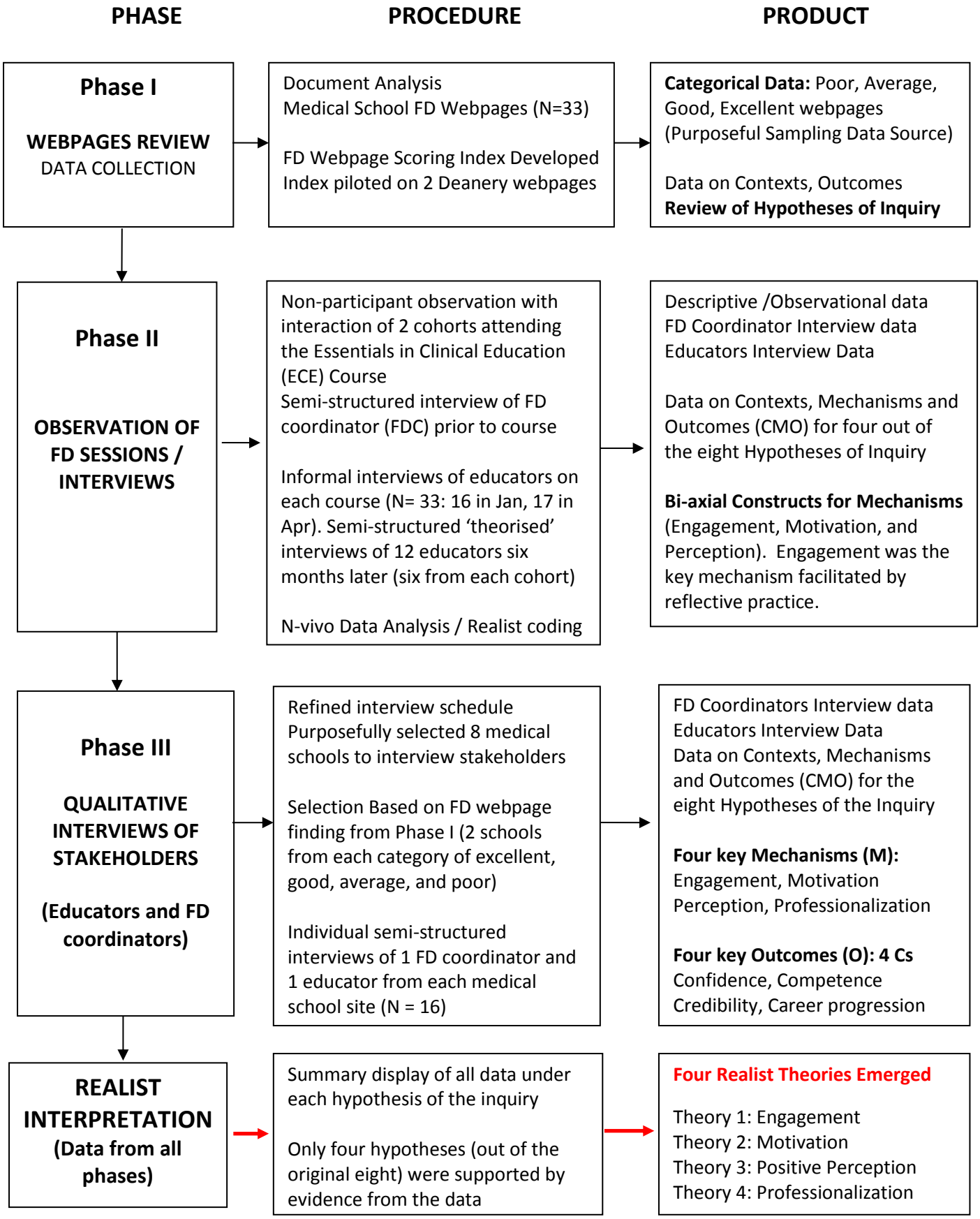
# CHAPTER 9
# DISCUSSION

## 9 Introduction

The purpose of this research study has been to evaluate the effectiveness of FD in UK medical schools through a realist approach and explain outcomes in terms of mechanisms acting in contexts. To do this, hypotheses were constructed (section 3.2.4). Phase 1 of the research process involved reviewing UK medical schools' webpages for the presence of FD activities, their contexts and expected outcomes. Phase II involved my attendance at a FD course with observations and interviews of educators to establish *CMO* connections in the short term (Chapter 7). Phase III considered longer term data from educators and FDC that supported, modified or invalidated the hypotheses (Chapter 8). A summary of how the data were collected and analysed in the three phases of the study is shown in Figure 9.1.

In this chapter, I explain how the findings from Chapters 6, 7 and 8 were synthesized to evaluate which hypotheses might be upheld and which might be unsupported. As a result of this process some hypotheses remained unaltered, some were changed, and some were discarded. This is followed by a discussion of my own views of the realist theories that emerged and the implications for FD. Next I review the use of the realist framework and the modifications I had to make to use it in this study. Finally, I consider data quality in terms of validity, reliability and researcher effect that could have compromised the final theories that emerged. The chapter ends with a summary of the key points.

## Fig 9.1: Diagram of the Phases



## 9.1 Data display

My purpose in collecting the data was to verify, modify or disprove the hypotheses hence I decided that the data should be presented as reflective summaries each headed by a hypothesis of the inquiry (from Table 3.2). Below each summary will be a table of the relevant hypothesis that pertained to it. I chose this approach after being influenced by the effectiveness of visual display of data in the work of Miles et al. (2014), where this is central to their approach to qualitative data analysis. When the data are presented in this orderly and systematic display it could be used for drawing conclusions.

To check on the validity of my judgement, once the summaries were completed, I referred again to the interview notes to ensure that my summary reflected the views of the stakeholders. When needed, I also referred back to the appendices and the full details of the interviews. Furthermore, in order to make sense of the available data, I used my skills as a clinician and a researcher and followed Clarke's (2004) advice that researchers should use their insight as well as information that has been collected in a scientific way,

> ***"Rigour is not everything. Saying only what you can say with a high degree of certainty is often less important and less useful than doing the best you can with the information available, and in the time available." (Clarke, 2004, p. 84)***

## 9.2 Review of the hypotheses

I have reviewed each hypothesis in turn and made clear the source of data. The alterations to the hypotheses came from my interpretation of the appropriate data. As

can be seen in the summaries below, for some hypotheses there were no data that pertained to contexts, mechanisms or outcomes from the phases.

### 9.2.1 Hypothesis 1 (Engagement)

There was extensive information that related to hypothesis 1 as there was general agreement with this hypothesis by all stakeholders. This developed as one of the most robust hypothesis as data emerged from the different aspects of the research project that supported the hypothesis. Furthermore, the elements and powers within the context, mechanism and outcome illustrated the complex layers within the hypothesis as evidenced from the whole research study. There was evidence from phase I regarding participatory approach as a contextual factor. Phase II was able to provide further descriptors of engagement using the bi-axial constructs (section 7.2.2) with a deeper understanding of engagement emerging as interactive and informative FD. This phase also provided evidence on engagement maintained by reflective practice and confidence being a key outcome for educators.

Phase III provided longer term evidence that corroborated the short term evidence from phase II. Both educators and FDC emphasised the importance of engagement for learning. Furthermore, there was robust evidence on reflective practice as facilitating engagement in the long term as well as increased confidence in teaching being the core outcome as discussed in Chapters 7 and 8. My interpretation of all the collected data was that of strong evidence and agreement with the context, mechanism and outcome of hypothesis 1 (Table 9.1). Moreover, the finding is in agreement with the literature on engagement being very important for deeper learning.

**Table 9.1: Hypothesis of the inquiry 1**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 1 | FD uses participatory approach and encourages reflective practice | + | Engagement generated by the FD being very interactive and informative using multimodal approach | = | Increased confidence in teaching. Empowerment to utilise teaching opportunities |

### 9.2.2 Hypothesis 2 (Motivation)

There was evidence from phase I (Table 6.3) for the contextual aspect of this hypothesis in combination with data from the literature review. Phase II provided evidence on the mechanism (individual motivation) and the outcome (improvement in instructional skills). Phase II findings revealed that the most important motivation was individual rather than altruistic, internal or external (section 7.3.1). Furthermore, the top three individual motivators were career development (improve CV, qualification), personal teaching interest (interest in teaching, opportunity to update), and personal development (to become more confident, opportunity to have some teacher training). These are important points for FD developers to note. It also answered Steinert et al's. (2006, p. 519) question following their systematic review of FD when they suggested that the question of motivation to attend FD activities remained unanswered and asked: "What motivates participation in FD?"

Phase III provided further support for hypothesis 2. On context, this phase showed that a FD course relevant or aligned to the needs of the educator is an important contextual factor for motivation as commented on by both educators and FDC. This phase further extended and clarified the outcome as educators described the competence that followed attendance at FD courses (section 8.2.3.2). Understandably in phase III, educators have had a longer time to assess and form a

view of their own teaching competence and performance compared with the shorter time period in phase II. As discussed in the previous chapter, competence was the second key outcome identified by the educators. Competence was conceptualised as combining the knowledge from attending the FD together with the expertise developed from experience and practice and applying it with the right attitude and values in their teaching. The hypothesis therefore remained unchanged in substance, although as discussed in Chapter 8 it was rewritten as hypothesis 2c to make the outcome clearer (Table 9.2).

**Table 9.2: Hypothesis of the inquiry 2 (Final version 2c)**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 2 | FD course aligned to the needs of the educator or FD relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Improvement in instructional skills and change in teaching approach |
| Hypothesis 2c | FD course aligned to the needs of the educator or FD relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Competence as shown by improvement in instructional skills and change in teaching approach |

### 9.2.3 Hypothesis 3 (Positive Perception)

Most of the evidence for this hypothesis came from the data presented in Chapters 7 and 8. In section 7.3.3, perception of FD by the educators was considered in the domains of usefulness and relevance on the bi-axial constructs. There was substantial evidence provided by the educators who attended the course on the usefulness and relevance to their own teaching. Phase III provided more detailed evidence on positive perception of FD. There was evidence on the contexts of access to FD and access to facilitators. FD coordinators provided details of the

various methods they adopted to provide access to FD. Furthermore, perception as a mechanism was conceptualised in terms of the value of teaching: the individual and institutional values of teaching enhanced by FD. At the individual level, the educators felt that learning from FD made them appreciate the importance of teaching; they felt more enthusiastic and interested in teaching and valued its contribution to their own professional practice and quality improvement. At the institutional level, educators felt that participation in FD had altered the institutional perception of their teaching. They felt they were better supported to attend other teaching courses, their ideas and contributions on teaching were considered more, and there was acknowledgment of the teaching component in their promotion. So although there were comments by few educators that their institutions could do more, overall majority felt that their institutions valued their teaching.

Phase III also provided evidence of a more extended and detailed outcome for hypothesis 3 as educators commented on credibility. Educators discussed credibility in terms of application of their knowledge and expertise to teach, in terms of the relationship with their students and colleagues, and in relation to how they are perceived by the educational establishment. My interpretation is that credibility (following on from confidence and competence) is the third important outcome for educators. Therefore, hypothesis 3 remained unchanged in meaning but was rewritten as hypothesis 3b (section 8.3.3.3) to show a more detailed outcome (Table 9.3).

**Table 9.3: Hypothesis of the inquiry 3 (Final version 3b)**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 3 | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Increased enthusiasm, interest and higher importance attached to teaching. |
| Hypothesis 3b | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Credibility as a teacher / increased enthusiasm, interest and higher importance attached to teaching. |

### 9.2.4 Hypothesis 4 (Feedback)

The data collected on this hypothesis were varied. Part of the difficulty might be because feedback could be viewed either as mechanism or outcome depending on the interpretation of the data. Feedback could be a mechanism if it was conceptualised as information provided by an agent (teacher, peer, parent) regarding aspects of one's performance or understanding (Hattie & Timperley, 2007). Hattie and Timperley (2007) suggested that feedback was one of the most powerful things that teachers can do to enhance achievement of their learners. However, they further explained that feedback could be differentially effective and for feedback to be a mechanism for learning, it needs to provide information specifically related to the task or process of learning that fills a gap between current and desired understandings. There must be a learning issue which the feedback addresses. Moreover, feedback is not necessarily a reinforcer, as it could be accepted, modified or rejected by the recipient. In my literature review, there were examples of feedback as mechanism; Marvel (1991) and Steinert et al. (2006) reported that individualised,

systematic and constructive feedback resulted in improved teaching performance (section 2.9).

In phase II, I considered feedback as a mechanism. During the course there was feedback from peers (during skill demonstration and presentation sessions), and from facilitators (to the small groups and individuals at times during tasks). Post-course, there was feedback from the peer observation of teaching practice and from the assessors to participants following submission of their portfolios. Admittedly, most of the feedback in phase II was generic (rather than individualised) and based on a framework of giving constructive and meaningful feedback to enhance learning. However, participants valued the feedback as it gave them a framework to practice with and improved their understanding. Furthermore, as discussed in section 7.4.2.4, feedback was one of the specific skills that some participants were deficient in; hence it could also be considered a context which motivated learners to learn and improve their feedback / instructional skill. My interpretation based on my contemporaneous observer record is that feedback on the course was mostly generic and focused on the practical aspect of how it should be done. Most educators did not strongly consider it a key mechanism that underpinned their learning during the FD as it was mostly generic.

In Phase III, most of the evidence in relation to feedback was as an outcome in terms of improved feedback skill as educators changed how they give feedback to others. Educators gave some examples of how they have altered their feedback in various scenarios such as in marking written papers, practical skill demonstration, simulations and clinical teaching. In fact, hypothesis 2 was rewritten as hypothesis

2b (section 8.3.2.5) with feedback added to the outcome as shown in Table 9.4. However, my overall interpretation is that while educators understood the importance of feedback, none of the educators or the FDC in phase III identified feedback as a possible mechanism for learning from the FD activity, hence I have not included hypothesis 4 (Table 9.4) in the final theories.

**Table 9.4: Hypothesis of the inquiry 4**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 2b | FD course providing training on feedback skill relevant to the need of the educators | + | Motivation to learn the necessary skill in areas of weaknesses | = | Improvement in feedback skills / change in educators behaviour e.g. in giving feedback to others |
| Hypothesis 4 | FD using an iterative cycle of training and continuous dialogue with stakeholders | + | Feedback to the educators during the FD and / or assignments | = | Improved teaching performance |

### 9.2.5 Hypothesis 5 (Collaboration)

Data collection for this hypothesis was limited. Even though the evidence from the literature review (Table 2.4) suggested collaboration as a possible mechanism with the sense of ownership and commitment to teaching as an outcome, there were no data from any of the phases to support this. This was probably because the current approach to the design of FD interventions does not involve joint collaboration between FDC and educators. However, as discussed in section 8.3, both stakeholders envisaged more local ownership of FD using a collaborative model based on needs assessment and local context so that educators have more ownership of how their FD is shaped and delivered. My interpretation is, that at the current time, hypothesis 5 (Table 9.5) is not supported but this area needs more

research as we have to think of tailoring FD content to the participants' needs and outcomes desired.

**Table 9.5: Hypothesis of the inquiry 5**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 5 | Sense of joint responsibility with shared needs in teaching | + | Collaboration between FD coordinators and educators in designing programmes / strategies | = | Sense of ownership, shared understanding and commitment to teaching |

### 9.2.6 Hypothesis 6 (Professionalization)

In the literature review, I noted that there have been many government initiatives recommending professionalization of medical school teaching (section 2.4.2) therefore it seemed logical to include this hypothesis. Findings from Phases I and III provided evidence on context indicating that a range of organisations and bodies were involved in target setting and suggesting strategies for medical education. There was also evidence in phase III that FDC strongly viewed this as an important mechanism even though there were fewer data from the educators themselves regarding professionalization as a mechanism. This could be due to the fact that at the time of their interviews, most educators were attending voluntarily rather than due to external pressure. Despite this, half of the educators in phase III talked about professionalization in terms of providing a sense of purpose to their learning, a sense of achievement upon completion as well as the recognition gained career wise. There was plenty of evidence from the educators on teaching career progression which was directly attributed to the FD attended confirming that professional status and recognition did bring about change. Hence, I have retained this hypothesis while

acknowledging that though the data on its mechanism might be limited, there was overall support for the hypothesis across the three phases of the research study.

**Table 9.6: Hypothesis of the inquiry 6**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 6 | Outside initiatives / external influences and demand to standardise medical educators teaching | + | Professionalization of teaching with qualification/ accreditation and standards by recognised bodies or authorities | = | Career progression in teaching / improved recognition as educators |

### 9.2.7 Hypothesis of the inquiry 7 (Time)

Most FDC did mention time constraints and competing pressures as a challenge for designing and delivering varieties of FD programmes. However, from the educators' point of view because of the method of sampling and the design of the study (all already attended FD activities), it was difficult to obtain data that would support the context and mechanism of this hypothesis. Hence, I have to accept that hypothesis 7 (Table 9.7) cannot be reviewed since the data from the study does not support or challenge the hypothesis.

**Table 9.7: Hypothesis of the inquiry 7**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 7 | Time available to attend FD / Time available to practice what was learnt | + | Regular attendance and participation at FD and updates | = | Improvement in instructional skills / improved student feedback or student evaluation of teachers |

### 9.2.8 Hypothesis of the inquiry 8 (Training)

If FDC are not properly trained for their role, then FD programmes would probably be poorly designed and educators would not benefit as much from them. There was no evidence to support this hypothesis. The FDC who were the stakeholders with expertise to comment on this, spoke freely about their background in teaching and the requisite qualification for their job. Most had considerable teaching experience prior to taking on their current role. They stated that the fact that less than half were from health related disciplines made no difference to their ability to design and develop teaching activities for medical school faculty. The educators stated that the FDC were knowledgeable with considerable expertise. In my view, this was a hypothesis where the responses underlined the advantages of a good researcher relationship where the stakeholders were able to freely give open and honest answers. It could be argued that this is a hypothesis that could be regarded as 'simply true' (Pawson, 2002c) and needed no further evaluation. However, my interpretation is that though the hypothesis remained unchanged in its meaning, it can be re-phrased positively as hypothesis 8a (Table 9.8) to reflect the information from the stakeholders.

**Table 9.8: Hypothesis of the inquiry 8**

| | Context | | Mechanism | | Outcome |
|---|---|---|---|---|---|
| Hypothesis 8 | FD coordinator with little or no training for the role or no time to develop the role | + | Poorly designed FD programmes on offer for educators | = | Negative outcomes for educators as no learning acquired |
| Hypothesis 8a | FD coordinators properly trained and their role in teaching clearly defined | + | Coordinators develop and offer well designed FD programmes on teaching | = | Positive outcomes for educators as good teaching skills acquired |

### 9.2.9 Summary of the theories

The four hypotheses (1, 2c, 3b and 6) supported by the available data are now rewritten as theories 1, 2, 3, and 4 as shown in Table 9.9 below. Bearing in mind that FD is a complex social intervention the end result, as Pawson (2013, p.112) stated, will be "partial knowledge about partial improvements we can make in the delivery and targeting of social interventions." My interpretation of this is that any or all of these mechanisms could be operating singly, jointly, simultaneously or cumulatively to produce any or all of the various outcomes depending on the individual context(s) of each educator. For example, while motivation and engagement seem to be very important in the short term, engagement and perception were more important in the longer term. The answer to Steinert et al's (2006, p. 519) second question: "What determines whether someone will take advantage of specific FD interventions at a particular time?" is to be found in the contexts and mechanisms of the theories below. Furthermore, it is important to remember that while programmes offer resources, whether they work depends on the choices of the subjects i.e. programmes work only if people choose to make them work (Pawson, 2006a). I will discuss each of the final theories in the next section with this in mind.

**Table 9.9: The four realist theories supported by the CMO findings**

| | Contexts | | Mechanisms | | Outcomes |
|---|---|---|---|---|---|
| **Theory 1 Engagement** | FD uses participatory approach and encourages reflective practice | + | Engagement generated by the facilitator being very interactive and informative using multimodal approach | = | Increased confidence in teaching. Empowerment to utilise teaching opportunities |
| **Theory 2 Motivation** | FD course aligned to the needs of the educator or FD relevant to the learners needs e.g. aspects of teaching they are deficient in | + | Motivation to learn the necessary skills in areas of weaknesses | = | Competence as shown by improvement in instructional skills and change in teaching approach |
| **Theory 3 Positive perception** | FD in a setting that facilitates the educators in their job (including access to FD activities and to FD coordinators) | + | Positive perception of the value and relevance of teaching in their job and in the organisation | = | Credibility as a teacher / increased enthusiasm, interest and higher importance attached to teaching. |
| **Theory 4 Professionalization** | Outside initiatives / external influences and demand to standardise medical educators teaching | + | Professionalization of teaching with qualification / accreditation and standards by recognised bodies or authorities | = | Career progression in teaching / improved recognition as educators |

## 9.3 Implications of the realist FD theories

In this section even though I discuss each of the theories individually, it is important to appreciate the interconnection between the four theories. For example while engagement is regarded as the bridge between motivation and learning (Hargreaves, 2006), it is also recognised that teachers can enhance engagement by building into the work, the interests and needs of the learners (i.e. motivation). Similarly, Swanson (2009) has commented on how values and perceptions influence engagement while Newswander and Borrego (2009) reported on engagement being enhanced by the granting of degrees and the provision of tenure. Perhaps what is

noticeable is that engagement is central and seems to connect all strands of the four theories.

### 9.3.1 Realist Theory 1: Engagement

There is no doubt that engagement is very important for deep learning to occur. Learners who are engaged learn at high levels and have a profound grasp of what they learn, retain what they learn, and can transfer what they learn to new contexts. The major premise of engagement theory is that learners must be engaged in their work for effective learning to occur (Miliszewska & Horwood, 2004). My finding on engagement in FD fully supports the literature on engagement and deep learning. Teachers can enhance the prospects of learners being engaged in the tasks and activities by carefully building into the teaching they provide those qualities that are most likely to appeal to the values, interests and needs of the students involved. An important finding in this study is the context that promoted FD engagement. So while a participatory, multimodal, interactive approach is important during the FD, long term engagement is facilitated by reflective practice which is important for FD developers to understand.

The concept of reflective practice as an intellectual activity has been recognised for a long time. Dewey (1933) saw reflection as a deliberative process with each thought carefully related to its predecessor and successor and which could profoundly influence one's experiences; he suggested that reflective thought was underpinned by evidence. Since Dewey, many authors have proposed varying definitions of reflection, which represent differing conceptions of the term (Box 9.1).

**Box 9.1: Definitions of reflection**

Dewey (1933, p.104): Active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions to which it tends

Schön (1983; 1987): The reflective practitioner is based on the notion that professionals need to learn to frame and reframe complex and ambiguous problems while engaged in practical situations as 'reflection-in-action' or after an activity as 'reflection-on-action', both forms resulting in reasoned judgments and ways to act.

Boud et al. (1985, p.19): Those intellectual and affective activities in which individuals engage to explore their experiences in order to lead to new understandings and appreciations

Kottkamp (1990, p.183): A cycle of paying deliberate, analytical attention to one's own actions in relation to intentions – as if from an external observer's perspective – for the purpose of expanding one's options and making decisions about improved ways of acting in the future, or in the midst of the action itself.

Killion and Todnem (1991, p.15): Expanded Schon's reflection model to include the concept of 'reflection-for-action' which involves thinking about future actions with an intention to improve or change one's practice. This type of reflection guides future action based on past thoughts and actions.

Hatton and Smith (1995, p.40): Deliberate thinking about action with a view to its improvement

Epstein (2008, p.1048): Reflection is promoted as one way of gaining access to perceptions and judgements that often escape our awareness, and thus may place us in a better position to change them. Reflection involves metacognition (thinking about thinking and feeling about feeling)

Whatever the nomenclature, self-awareness, critical analysis and the development of a new perspective are fundamental to the process of reflection. The important thing is to identify the diverse ways in which reflective practice can enhance the development of faculty members. Therein lies the problem as much of the literature has focused on how the reflective capacity of learners could be developed but rather fewer publication providing evidence that encouraging learners to reflect improves their resultant actions. To some extent this remains an assumption, albeit one that is underpinned by a number of seminal pieces of literature. One of its key proponents is Moon (1999). She argues that by reflecting, learners are able to 'upgrade' their learning to even higher levels after the original time of learning (Moon, 1999, p.147). My study has provided further support for the context of reflective practice leading to engagement and deeper learning. FD developers and coordinators have to

incorporate and actively encourage reflective practice in their FD design to promote learning. Sorinola and Thistlethwaite (2013) in a systematic review have shown the impact of context on the success of educational initiatives. This reflective process of self-assessment and critical analysis of teaching allows for integration of theoretical concepts into practice, greater learning, enhanced critical thinking and judgement.

The outcome from this engagement theory is also quite important. Confidence in teaching was expressed in various ways such as in delivery of teaching, breaking down complex teaching activities into understandable components, trying out new instructional techniques, dialogue with students and peers, and examination of teaching practices. Confidence in teaching is an important springboard for educators to develop the other outcomes, hence FD needs to facilitate this outcome.

### 9.3.2 Realist Theory 2: Motivation

The Latin root of the word ‘motivation’ means ‘to move’ (Eccles & Wigfield, 2002), hence theory 2 on motivation does help us understand the role played by expectancies, needs and values (all related to motivation) in helping educators ‘move’ to participate in FD and learn from FD activities. Motivation theories have evolved over time from the early theories on intrinsic and extrinsic motivation (Knowles, 1978; Knowles et al, 1998) to other motivation theories that suggest that to construe the mental event of motivation as a simple dichotomous internal or external phenomenon is to deny the very complexity of the human mind (Box 9.2). Although design, objectives, and measured outcomes have varied between studies, most have reported a positive correlation with motivation influencing learning, study behaviour, academic performance and success (Wilkinson et al, 2007; Sobral,

2004). So my theory in general does provide additional support for the literature on the importance of motivation to learning.

**Box 9.2: Motivation Theories**

Attribution Theory: Perceived causes of an event or outcome shape behaviour, affect and cognition (Weiner, 1974).

Expectancy-Value Theory: Positive expectations and high value placed on task or outcome enhances motivation (Vroom, 1964; Atkinson, 1966).

Goal Theory: All human actions and behaviour are motivated by a goal; it examines the reasons why individuals engage in a particular behaviour or pursue a particular goal (Pintrich, 2000).

Self-Theories of motivation: (Bandura, 1997; Reeve et al, 2004; Dweck & Molden, 2005).

Specifically for FD though, the literature on motivation is scant: apart from Steinert et al. (2010) no one has looked at what motivates educators to participate in FD initiatives. Most programme developers have extrapolated form the broad field of continuing medical education (CME) where several authors have examined what motivates doctors to attend formal events (Harrison & Hogg, 2003). Motivators identified in those studies include: maintenance of professional competence; acquisition of new knowledge and skills; improvement of understanding of key concepts; elimination of clinical deficiencies, and reassurance that one is 'doing it right' (McLeod & McLeod, 2004). So the issue was whether these same motivators were involved in FD. Steinert et al. (2010) used a focus group to explore the motivation of teachers for attending a FD course and reported the following five factors: learning and self-improvement are valued; FD topics has to be relevant to teachers' needs; FD is seen as enabling personal and professional growth; the opportunity to network with colleagues is appreciated, and initial positive experiences promote on-going involvement.

So though my findings on motivation are similar to that of Steinert et al. (2010), there are additional findings from the realist approach adopted. The context of relevance of FD to the needs of the learners is important. I have shown that individualistic factors (career development, personal teaching interest and personal development) were the most important motivators for the educators. So while I conceptualise motivation as multifaceted, multidimensional and dynamic with educators able to move between different types of motivational state depending on the situation, it is crucial for FD developers to understand the highly individualistic drivers in FD. Furthermore, I have been able to show the link between motivation to learn and the competence developed by educators with improvement in instructional skills and changes in teaching behaviour. The importance of this motivation theory is that it contributes to a more complex understanding of the nature of motivation in FD. The current paradigm of motivation in FD has to broaden to incorporate these findings. If we continue to create FD programmes without improving our understanding of the nature of educators' motivation we might be missing an opportunity to significantly enrich their training. Motivation should be a measure of the quality of FD initiatives because if we truly value the learner who is engaged, who participates, who learns deeply, then we value a learner who is highly motivated. What is the development of a 'life-long-learner' in FD if not a learner who is individually motivated?

To summarise, realist theory 2 on motivation can help FD developers plan how to trigger educators' expectancies in FD by addressing learners' needs for competence and autonomy. Each need fulfilled on its own promotes motivation, however, fulfilling both needs at once creates a synergistic effect. The need for competence is fulfilled by making FD relevant to learners' needs, providing optimal challenge and positive

performance feedback, and the need for autonomy by providing choice in the types of FD provided and opportunity for self-direction. It would therefore seem worthwhile to explore motivation and its contexts in FD to identify the relevant needs and expectancy beliefs at play when faculty members choose to participate in FD activities.

### 9.3.3 Realist Theory 3: Positive Perception

Perception as explained by philosophers is a complex neural process shaped and affected by a variety of factors. According to Weiten (2013), humans have a perceptual expectancy i.e. a predisposition to perceive things in a certain way which is an example of how perception can be shaped by 'top-down' processes such as drives, motivations and expectations (Coon & Mitterer, 2009). Clark (2011) further explained that perception, is not simply a 'bottom-up' process where minute details are put together to form larger wholes, instead there is a great deal of feedback between perception and expectation (as perceptual experiences often shape our beliefs, but those perceptions are based on existing beliefs). Thus, there is a link between perception and motivation when viewed through the expectancy-value theory as described by Fishbein (1975). According to this theoretical framework, we orient ourselves to the world according to our expectations and values (Palmgreen & Rayburn, 1985). Eccles and Wigfield (2002) have defined expectancies as 'beliefs about how one will do on certain tasks or activities' and values as the 'incentives or reasons for doing the activity'. In diverse ways, these expectancies may be at play in FD. For example, if teachers believe that FD can enable personal and professional growth and that it is relevant to their needs, they may be more likely to participate

and learn as reported in phases II and III findings. In addition, their appreciation of teaching and self-improvement may function as motivators with the end result of increased enthusiasm / importance attached to teaching and improved credibility (outcome).

Interestingly, most of the work related to perception and expectancy-value theory has been conducted with elementary and high school children. My study contributes to exploring this theoretical framework in the context of higher education by identifying the types of expectancy and values at play when faculty members choose to participate in FD activities. Most educators reported a positive perception of FD and valued FD at both individual and institutional levels. Therefore, FD developers have to understand the importance of designing programmes that adds value at both individual and institutional levels.

### 9.3.4 Realist Theory 4: Professionalization of medical education

The issue of professionalization of medical education has received increased attention over the past several years in response to growing pressure of the need for standards amongst medical teachers and the view that teaching is a core professional activity that cannot be left to chance, aptitude or inclination (Purcell & Lloyd-Jones, 2003; Eitel et al, 2000). Furthermore, the concern remains that most medical school faculty members have had little or no training in *how* to teach (MacDougall & Drummond, 2005). The first challenge though is understanding the terminology. The literature is replete with many attempts at definition, but the semantics of professionalism in medicine (a state) and professionalization of medical education (a process) obfuscate more than they clarify as some authors use the

terms interchangeably while others conflate them, for example, Cruess et al. (2004) (Box 9.3). Therefore, for clarity and to make the distinction, I used Hafferty's (2008) definition of professionalism (Box 9.3) and considered professionalization of medical education as an attribute generating process that applies to individuals as well as institutions. It is a process to attain and maintain professionality (compliance with agreed standards) as explained by Eitel et al. (2000).

**Box 9.3: Definitions of Professionalism**

*Professionalism in medicine* (Hafferty, 2008, p. 21)
The altruistic, ethical and values aspects that resides in the interface between the possession of specialised knowledge and a commitment to use that knowledge for the betterment of others.

*Professionalism in medical education* (Cruess et al, 2004, p. 74)
Profession is an occupation whose core element is work based upon the mastery of a complex body of knowledge and skills. It is a vocation in which knowledge of some department of science or learning or the practice of an art founded upon it is used in the service of others. Its members are governed by codes of ethics and profess a commitment to competence, integrity and morality, altruism, and the promotion of the public good within their domain. These commitments form the basis of a social contract between a profession and society, which in return grants the profession a monopoly over the use of its knowledge base, the right to considerable autonomy in practice and the privilege of self-regulation. Professions and their members are accountable to those served and to society.

So for medical educators having the requisite knowledge in teaching by attending FD to gain teaching qualification and accreditation by recognised bodies is an important component of professionalization. It certainly was a driver for learning during the FD as educators became increasingly aware of the need for certification and accreditation for career progression and credibility as educators. This was expressed in various ways including job opportunities, promotions and leadership roles. My study supports the literature finding that standards can help to raise the status of medical educators and improve the quality of FD provision (Wojtczak & Schwarz, 2000; Eitel et al, 2000). Other authors have also suggested that an advanced degree in medical education offers essential grounding in educational theory and practice

and provides a foundation for educational research, scholarship and leadership (Cohen et al, 2005).

While several bodies / authorities, nationally and internationally (GMC, WHO, WFME), have advocated the need to certify medical educators to ensure standards and improve the quality of medical teaching, some authors have sounded a note of caution about disenfranchising keen and committed educators (Eitel et al, 2000; Purcell & Lloyd-Jones, 2003). In my view, there lies the importance of understanding the context for this mechanism in that a balance has to be struck between the external influences, demands, recommendations (including resources needed) and the need for certification and accreditation. The realist maxim 'what works for whom' becomes important at both individual and institutional levels.

In summary, I believe that professionalization is multi-dimensional and while certification / recertification / accreditation is one aspect, another aspect (outside the scope of this study) could be implementation of a faculty evaluation programme integrating data from students, peers and administrators to provide meaningful evaluative information for faculty. After so much debate and publication, one would expect that the definition and important attributes of professionalization would be well codified by now, but the evidence in the literature suggests not. Hopefully my study would help in expanding the understanding of this area.

## 9.4 Review of realist evaluation design and modifications

In this section, I review the use of realist evaluation as a framework for educational research. My use of Pawson and Tilley's (1997) model in education research was

distinctive since it had been used, previously, mainly in the field of crime prevention. To explain my adaption of realist evaluation for medical education, each stage of the research process from the design to data collection and interpretation (Figure 9.1) is considered with a discussion of the development and refinement needed at each stage to modify the model.

### 9.4.1 Realist evaluation design

In Chapter 3, I discussed Pawson and Tilley's (1997) conception of mechanisms acting in contexts to produce outcomes which gave me a clear underpinning principle for the research study. I also used their diagram of the research cycle (Figure 3.2) to understand the realist design process. However, as discussed in section 3.2.1, I made a clear distinction between 'hypotheses' of the inquiry and the final emergent 'theories'. Furthermore, I used Maxwell's (2013) interactive research design model (Figure 3.3) to complement Pawson and Tilley (1997) idea and I developed an integrated, interconnected, flexible and iterative design model for FD in medical education.

I did consider the alternative model of programme evaluation used by Pawson (2002a) called the 'theories-of-change' model whereby the programme is conceptualised as a theory and the programme theory is presented as a chain of stages as it passes through initial activities, intermediate and long term outcomes. As described by Pawson (2002a), the idea is to collect data that will identify flows and blockages at each stage in the chain.

> *"The evaluation consists of putting a microscope to each stage, making process observations to see if the theories conform to actuality."*
> *(Pawson, 2002a, p. 473)*

Just as in realist evaluation, the theories-of-change model is based in the philosophy of realism and seeks to identify contexts that will facilitate or prevent outcomes. However, the theories-of-change model was developed for social programmes with long and complex implementation chains or, as Pawson (2006a) further described, it can be used for systematic reviews of evidence-based policy where researchers consider the results of previous inquiries. It would seem, therefore, that this model does have a place in educational research for the examination of educational policy, but at a central or governmental level rather than by individual practitioners. In my opinion, most educationists are interested in the effectiveness of discrete programmes such as FD. Therefore, the model of realist evaluation described by Pawson and Tilley (1997), and adapted with ideas from Maxwell (2013), was more apposite for my FD study.

### 9.4.2 Developing the hypotheses of the inquiry

In section 3.2.4 on developing the hypotheses of the inquiry, I followed Pawson's (2003b) guidance that the researcher should select those hypotheses that he or she thinks are vital to the effectiveness of the programme. Therefore, I selected hypotheses which in my judgement, reflected the main structures of the FD and then set out to collect data that would support, modify or invalidate them. However, it was not until phases II and III that I realised that the task of data collection for so many hypotheses was very challenging within the constraints of the time and resources of a single researcher.

The issue here is not in the design of realist research but in the interpretation of it. One could argue it would have been more practical to have selected about four or five hypotheses. These would still have needed to include the hypotheses, which were important for the effectiveness of the programme, but exclude some of the more obvious hypotheses. For example, hypothesis 7 on time (Table 9.7), which was about having time to attend FD and to practice what was learnt. Perhaps time is an obvious context for FD programmes and there was evidence from the literature review (Chapter 2) on the importance of time in FD interventions. Therefore, hypothesis 7 could have been excluded from the study and the same could be said for hypothesis 8 (training). On the other hand, I felt it was appropriate to include hypothesis 5 (collaboration) even though there was limited evidence to support it. My view is that realist research is about exploring the layers of a social programme and the effect of emergence when elements in the programme combine. Therefore it is important to include more complex theories, which can then be discussed with the stakeholders.

### 9.4.3 Data collection in realist evaluation

In realist research, the object of study defines the data collection method. Throughout this study I explained my choice of research methods in Phases I, II and III and followed them carefully (Figure 9.1). The framework described by Pawson and Tilley (1997) helped in clarifying the choice of method. They suggested three questions for the researcher: What do I need to know? Who can give me this information? And, how will I collect it? The framework guided me on the appropriate

research methods, and the stakeholders to consider. However, I did modify each stage of the research process to suit my study.

Pawson and Tilley (1997) discussed the kind of knowledge that different stakeholders might have about a programme. They suggested that participants could be sensitive to mechanisms but have less knowledge about contexts and outcome patterns. This method of allocating expertise (strictly to one group of stakeholders) was not very helpful in an educational study as I found both the educators and the FDC were knowledgeable on the various aspects of the *CMO* configuration and were able to offer some explanations on the various hypotheses.

What was more useful was the approach termed ‘different but complementary world views’ by Pawson and Tilley (1997, p.162). All the stakeholders, whether they were educators or FDC, were able to offer explanations on the *CMO* configuration of the hypotheses but from different knowledge bases and hence the viewpoint was qualitatively different. An example is Theory 4 on professionalization (Table 9.9) which as already discussed, some educators had limited knowledge or awareness of hence were unable to talk with authority about government or outside initiatives in relation to FD or how those initiatives are incorporated into medical school FD. Therefore, where data were missing or limited, this could be accounted for, in part, by the design of the study. Another possible reason for the inadequacy of data under certain theories could be the way that I chose to interpret the design of realist interviews. In phases II and III, I decided, in the teacher-learner phase of the interview, not to specify the hypotheses in detail but to explain the overall conceptual idea of the study so that participants had a clear understanding of the purpose of the

research and then used relatively broad questions. I chose this design since I believed (as explained in section 1.2.2) that the research project should be 'with' rather than 'on' the stakeholders since they had much to contribute to the process of the research. However, it was possible that I could have been more directional and not compromised the accuracy of the data from the interviewees and yet been able to guide the discussion towards more of the hypotheses. It is possible that a more careful design could have led to a wider spread of data collection. Therefore, although the final theories were a honest attempt to make sense of the data and of the effectiveness of FD, some of those theories involved my interpretation as a researcher.

To summarise, an important consideration is whether the final theories could be induced from the data that was collected bearing in mind that data to uphold, modify or disprove theories could sometimes be limited. Therefore, in realist research, it is important the researcher chooses sufficiently diverse groups, so that complementary knowledge on different aspects of all of the theories is ensured.

### 9.4.4 Data analysis in realist evaluation

This was perhaps the most challenging aspect of the realist evaluation of FD. In section 5.5, I discussed the analytic strategy for the mixed method data (quantitative and qualitative) collected. Pawson and Tilley's (1997) framework did not provide a clear method for analysing the data collected from the stakeholders: whether they used only their insight to draw conclusions from the data or whether some systematic analysis of the data was done. For my study, I rejected the former approach since I was concerned that it would be difficult to justify the validity of the

conclusions based only on insight and, instead, I used the method of qualitative analysis suggested by Miles et al. (2014). This modification was necessary, as I needed a method for aggregating the information from the different stakeholders. I used Miles et al. (2014) model to display and summarise the data under each hypothesis as discussed above. In addition, to decide on the final theories of FD, I used the summarised data (as guidance) and my judgement as an educator and a researcher. This idea, suggested by Clarke (2004) that the researcher should review all the data from the research inquiry but the final conclusions should also be guided by insight and understanding, was my final modification to the analytic template of the realist framework.

Based on the above, it is apparent and quite important that researchers need to modify the realist model based on the object of their study and adapt it for their own use. I will discuss this further in the next chapter under proposed recommendations for realist researchers. However, in the next section I will discuss the issue of data quality / trustworthiness and the strategies that I adopted.

## 9.5 Quality of data

It is important to consider the issue of data quality with regards to answering the original research question and in verifying that the final findings were authentic. I considered all aspects of the research process, which could have compromised the final theories. While I admit that it is an unrealistic goal to be 100% authentic (that is the optimism of perfection), as Wolcott (1990, p. 121) suggested regarding validity, it is also important “not to get it all wrong.” Various authors have stressed the importance of assessing the quality and trustworthiness of qualitative data using

various strategies (Robson, 2011; Miles et al, 2014; Cohen et al, 2011; Maxwell, 2013). I have adopted my data quality evaluation strategy from these sources and considered the following three areas: validity, reliability and researcher effect.

### 9.5.1 Validity

Validity refers to the issue of whether an indicator (or a set of indicators) that is devised to gauge a concept really measures that concept. Validity as defined by Hammersley (1990) is to do with accuracy.

> ***"By validity, I mean … the extent to which an account accurately represents the social phenomena to which it refers." (Hammersley, 1990, p. 57)***

Assessing the validity of the data for this research needed care since the study has both quantitative and qualitative components carried out within a realist framework. I therefore chose to follow the approaches of Miles et al. (2014), Cohen et al. (2011) and Bryman (2012) by considering internal and external validity since these applies in both quantitative and qualitative research.

#### 9.5.1.1 Internal validity

Internal validity seeks to demonstrate that the data collected accurately describes the phenomena being researched and to show there is a good match between researchers' data and the theoretical ideas they have developed, in other words, a check to see if the findings make sense and are credible. I explored internal validity under four criteria: representative data, respondent validation, triangulation / complementarity, and data analysis.

### *a. Representative data*

The first issue that I considered was about how representative is the data. In phase I (FD webpage analysis), I developed each criterion to evaluate a different component of the webpage. I asked two colleagues to independently comment on whether each of the criteria reflected the concept to which it was applied. Finally I tested the webpage criteria on the two postgraduate deanery websites before applying it to the medical school webpages as described in Chapter 5. I adopted a similar approach for the observed engagement scale in phase II. Once the scale was developed, I asked two colleagues to independently review the scale and I also piloted the scale on a different course.

Furthermore to improve validity, I re-assessed the methods used for data collection as suggested by Robson (2011), with the assumption that if the data collected were of a high quality, then this would contribute to validity. In the first part of phase II (observation and interview during the course), data were collected from as many participants as possible and the use of two cohorts improved the representativeness of the sample. For the six months follow-up interview I chose, by purposive sampling, educators with high, medium and low engagement scores with the intention of stimulating discussion on the various hypotheses.

In phase III, the medical schools (and therefore the FDC) were chosen by purposive sampling based on the FD webpage result and the chosen schools were spread across England, Scotland and Wales (Chapter 5). However, the educator lists were provided by the FDC and their choice of educators included on the list could have been influenced, perhaps subconsciously, by their beliefs in the positive outcomes of

FD so that only those educators they perceived to be of similar orientation were included. However, I did not think this was relevant as most of the educators listed were not personally known to the FDC and the FDC had no idea which educator I was going to choose from the list.

Data could also be non-representative if participants were biased or dishonest in what they said. I was concerned that the FDC could be influenced by their positive beliefs but, following their interviews, I was left with the impression of a very thoughtful and insightful group of people who were able to give a rich and detailed picture of FD and were able to discuss and criticise FD where necessary. For example, when discussing hypothesis 6 (professionalization) the FDC noted that sometimes FD could be viewed as a tick box exercise.

Apart from representativeness of the data, Cohen et al. (2011) suggested further criteria for assessing the validity of data. I attempted to ensure the *credibility* of the data by helping interviewees to understand the purposes of the research and by supporting them to ensure that they had the confidence to express their views. I also checked their responses to the various hypotheses with them. As detailed in Chapters 6, 7 and 8, *the kind and amount of evidence* that I collected was reasonably sufficient for the level of theorising even though as admitted above there was insufficient data for a few hypotheses.

### *b. Respondent validation*

Also known as member validation, this is a process whereby the researcher provides the participants with an account of the findings and solicits feedback (Maxwell, 2013;

Bryman, 2012). I sent my observation and interview data to individual participants in phase II and the interview data to phase III participants. This was to avoid misinterpreting the meaning of what participants said and did as well as the perspectives they had on FD. It was also an important way for me to identify my own biases and misunderstandings of what I observed or heard. I was aware that, as reported in the literature, respondent validation is not without potential practical difficulties. It may on the one hand lead to defensive reaction (and even censorship), but on the other there may be reluctance to be critical especially if there is mutual regard in the researcher-participant relationship (Robson, 2011). However, this was not an issue from my respondents in either phase II or III, and more specifically with my observation data in phase II where the risk of misinterpretation of intent was high. A possible explanation for this might be due to the fact that I was able to observe respondents over the three day period and informally interviewed them in a relaxed fashion (i.e. during breaks) to understand their thoughts and feelings on FD.

### *c. Triangulation and Complementarity*

Bryman (2012) defined triangulation as the use of more than one method or source of data in the study of a social phenomenon so that findings could be cross-checked. Maxwell (2013) expanded on this and said triangulation is the use of different methods as a check on one another, seeing if methods with different strengths and limitations all support the same conclusion (in this case the hypotheses). The point is to recognise the fallibility of any method or data and to triangulate in terms of validity threats. As explained in section 4.2, I used a variety of data collection methods (documents, observations and interviews) with different strengths and limitations. For

example, the strength of the FD observations was that it provided me with a direct and powerful way of learning about educators' behaviour and the context in which this occurred, while interviews of educators during the course provided additional information and a check on the accuracy of the observation. Interviews of educators in phase III allowed me to gain their views / descriptions of the FD they had attended in the past. As Weiss (1994) said,

> ***"Interviewing gives us access to the observations of others. Through interviewing we can learn about places we have not been and could not go and about settings in which we have not lived." (Weiss, 1994, p. 1)***

The second purpose for using multiple data collection methods was to gain information about different aspects of my research. For example, reviewing the FD webpage of a medical school and then interviewing the FDC and educator from the same medical school provided information on different aspects / perspectives on FD in that medical school. This is what Greene (2007) called *complementarity* where different methods are used to broaden the range of aspects of phenomena that are being addressed.

### *d. Data analysis*

During data analysis, I was also keen to ensure validity. In phase II, I compared my coding of the interview data with those of my three supervisors (in four randomly selected interviews) to check on the accuracy and coding agreement. Furthermore, during the development of the constructs for the mechanisms described in Chapter 7, I went through an iterative process with my supervisors and checked that the constructs were derived from and reflected participants' data. In phase III, I compared my coding of the first three interviews with that of an independent

qualitative data analyst to check agreement on codes and categories. I found the themes and categories were quite similar with similar data coded under context, mechanisms and outcomes.

#### 9.5.1.2 External validity

External validity refers to the degree to which the results can be generalised to a wider population or transferred to other milieu (Bryman, 2012). However, as Maxwell (2013) argued, generalisibility in realist research is not based on explicit sampling of some defined population to which the results can be extended, but on the development of a *theory* of the processes operating in the phenomenon studied, one that may well operate in other cases, but that could produce different outcomes in different circumstances. This is also sometimes referred to as analytic or theoretical generalisation (Robson, 2011). Hence, my interpretation for a realist study where the purpose was not to look for universal truth but rather for what works (M) for whom (O) in what circumstances (C), was to develop hypotheses that were framed in contexts, mechanisms and outcomes to explain aspects of FD in UK medical schools (Table 3.2). As noted by Pawson (2002c), social programmes are multifaceted phenomena which work in only limited conditions. This study, within the limits of the data gathered, had identified FD successes and recognised the conditions that facilitated those outcomes, thereby providing information for FD developers to identify the conditions needed for programme efficiency. In the postmodern conceptions of social sciences, the goal of global generalisation is replaced by a transferability of knowledge from one situation to another taking into account the context. As noted by Pawson (2003b),

*"We learn the transferable lessons about programme theories rather than about the programmes per se." (Pawson, 2003b, p. 479)*

### 9.5.2 Reliability

Reliability is essentially the degree to which a study can be replicated to achieve the same result (Bryman, 2012). It is a synonym for consistency, replicability and dependability (Cohen et al, 2011). This is a difficult criterion to meet in qualitative research since, as suggested by Bryman (2012), it is impossible to 'freeze' a social setting and the circumstances of an initial study to make it replicable in the traditional sense in which the term is usually employed. However, although the strength of qualitative research is its use of social interactions and the uniqueness of the data, this does not remove the need for qualitative researchers to strive for replication in the design and execution of their research. Various strategies have been suggested by various authors to control for reliability such as using a highly structured interview, standardised questions, and sequence for each respondent (Silverman, 2006). However, I rejected this design as I wanted an interview that would allow the interviewees to demonstrate their unique view of the world – in this case FD (Cohen et al, 2011). Instead, I controlled for reliability by pretesting the interview schedules, detailed how the interviews were conducted and how the data were coded and summarised (sections 4.2.3 and 5.5).

In qualitative research, strict replication is contentious as highlighted by Cohen et al. (2011) when they suggested that two researchers studying a single setting could come up with very different findings, but both sets of findings could be reliable and trustworthy. Therefore, although a researcher could set out in detail what he or she did, another researcher, following the same instructions, could collect quite different

data or provide different interpretation. This vicissitudes of interpretations was summed up by Kvale and Brinkmann (2009),

> ***"In interviewing, there might be as many different interpretations of the qualitative data as there are researchers." (Kvale and Brinkmann, 2009, p. 208)***

Hence, my interpretation of reliability in this FD study was to follow Bogdan and Biklen's (2007) suggestion and regard reliability as a fit between what researchers record as data and what actually occurred in the setting that was being researched i.e. a degree of accuracy and comprehensiveness of coverage. Following this construct, it could be said that the process of data collection and data analysis used in this study achieved a degree of reliability.

### 9.5.3 Researcher effect

Throughout this research study, I have been aware of my belief that FD could be an effective way to address the teaching needs of medical educators as discussed in section 1.2.2. I was cognisant that data collection and summary depends greatly on the judgement of the researcher and although my belief and experience with postgraduate FD motivated me to initiate the research, I took great care not to let it influence my judgment. Researcher 'bias' or subjectivity is an important issue and realist researchers need to account for the actual beliefs, values and dispositions that they bring to the study (Maxwell, 2012). Therefore, I was vigilant and careful during data collection (for example, I avoided leading questions, recorded observations accurately, used respondent validation and used purposive sampling). To summarise the data, I used my experience of note-taking and reviewing, then carefully read and

coded the data before placing them under the relevant hypotheses to ensure that the final displays (Tables 9.1 – 9.8) were representative of the data.

Reflexivity refers to the fact that researchers are part of the social world that they study and must therefore understand how they influence and are influenced by this world (Hammersley & Atkinson, 2007). Hammersley and Atkinson further said,

> ***"The researcher is the research instrument par excellence. The fact that behaviour and attitudes are not stable across contexts and that the researcher may play a part in shaping the context becomes central to the analysis …. The theories we develop to explain the behaviour of the people we study should also, where relevant, be applied to our own activities as researchers." (Hammersley and Atkinson, 2007, p. 21).***

This notion of researchers holding themselves up to the light was described as the 'looking glass self' (Cohen et al, 2011). I used my reflective memo written continuously during the research process to examine my feelings, assumptions and values. For example, I noted in my reflective notes that once the first session on day one of the course observation got underway, the participants hardly seemed to glance in my direction. My presence and effect on them ('reactivity') was no longer that obvious to me. I wrote that this could be because they were just immersed in the session or it could be that having informed them about my research prior to the course, supplied all the relevant paperwork, personally met and introduced myself to them during the registration period at the start of the course, they were quite content to get on with the course. In addition, having experienced the course with the participants, I had my own views of what worked for me during the FD in terms of engagement, motivation and perception but I had to set this aside during the analysis. So for me, although I had always been closely associated with a positive

belief in the FD, I was aware of that belief at all times and tried to ensure that it did not influence the way I conducted the research inquiry.

Another important consideration of the final theories is that there were no major changes and the final theories were similar in substance to the original hypotheses (see Tables 3.4 and 9.9). It could be argued that I am a researcher who was biased in favour of FD, developed hypotheses to support it and then set out to collect data that would uphold the hypotheses. However, as discussed above, I was aware of my positive feelings about FD throughout the study and endeavoured to control any confirmatory bias. One possible explanation for the constancy of the theories could be in the way that they were developed. I constructed the theories following a detailed literature review and from my own postgraduate experience / understanding of the main structures of FD (described in Chapter 1). Hence it could be argued that the hypotheses were written with some insight and, thus, the data collected generally supported them. Another possible explanation of why the theories changed little in substance could be attributed, at least in part to the stakeholders being positive about the structure of FD. They did not criticise the mechanisms of FD or offered explanations that involved a change in the mechanisms.

## 9.6 Summary

I return to my research question, which was at the top of the research cycle (Figure 3.2) to see if the study has been able to answer my question,

***Does evidence provided by the stakeholders suggest in what circumstances (C) Faculty Development Programme (M) can be effective for training medical educators and lead to good educational outcomes (O)***

The realist approach did answer the question and identified the four theories of FD (engagement, motivation, positive perception and professionalization). Realism assumes that each time a mechanism operates it can meet with success or failure depending on the context in which it operates. As the researcher, I had to analyse the workings of FD programmes to discover the contexts that produced successful outcomes and those that induced failure. However, the problem with such an approach is that social programmes are complex and operates at different levels, so collecting and analysing evidence was a challenging task. Despite this, the scope of the data collection was sufficient to indicate how certain aspects of FD were working and provided explanations for the outcomes. Moreover, the study has demonstrated the effectiveness of using a realist approach in medical education research by identifying the contexts in which FD programmes operates effectively.

This was an innovative study as the realist approach has not been used to a great extent in medical education. It has demonstrated the effectiveness of FD programmes in UK medical schools and the importance of contexts, which led to successful FD outcomes (confidence, competence, credibility and career progression). Moreover, this study has made a timely contribution to knowledge by providing policy makers with a template for evaluating FD practice within the UK medical school environment. This will be important especially with the recent GMC recommendation on undergraduate teaching.

Finally, as for any complex social programme, there are infinite numbers of hypotheses that could be explored in FD, but as a single researcher I had to decide and explore hypotheses that seemed relevant. Therefore the final theories had to be viewed as nested in a complex, dynamic, emergent world of FD. The realist approach was able to embrace this concept of complexity in social programmes by accepting that there is no 'absolute' truth (Pawson, 2013). In Chapter 10, I will consider and provide recommendations on FD and for the future use of the realist framework in medical education research.

# Chapter 10
# Conclusion and Recommendations

CONCLUSIONS AND RECOMMENDATIONS ON FD IN MEDICAL SCHOOLS AND THE USE OF REALIST EVALUATION

## 10 Introduction

In the previous chapter, I discussed the final four theories that emerged from this evaluation of FD for medical educators and included a detailed consideration of the validity of those theories. In this chapter, I will focus on the innovative use of realist evaluation in medical education FD; I consider this in two strands. The first strand is a reflection on the realist theories of FD from the standpoints of university policy makers, FD developers, and educators. This is followed by my recommendations for FD in UK medical schools. The second strand is a reflection on the realist framework itself with my recommendations on the future use of the model by other medical education researchers. I then discuss my personal reflections as a realist researcher during the PhD journey. In the final part of the chapter, I state my concluding message about FD in medical education.

### 10.1 Reflections on FD Theories

In this section, I discuss the implications of the four theories on FD evaluation for the different groups of stakeholders (university / medical school policy makers, FD developers and medical educators). I focus the discussion on the implication of these theories for FD practice in medical education. Nonetheless, FD need no longer be limited to teaching as there are other areas such as administration, management,

leadership and technology, which faculty seek training in hence a wider application of the theories will also be considered. A summary of the recommendations for FD is provided in Box 10.1.

### 10.1.1 University / Policy makers

How does this realist study help university and medical school policy makers? This study showed that FD in teaching needs to be valued at institutional level with provision of excellent webpages, adequate resources / funding, and FD should be made contractual for educators (both substantive and honorary). It is well known that while the mission statements of most medical schools advertise teaching as a priority, it is research that is usually prioritised (Clark et al, 2004; Steinert et al, 2005; Hitchcock et al, 1993). However, this trend will now have to change as this study serves as a timely contribution to the evidence and body of knowledge to stimulate this development (bearing in mind that the only UK wide study on FD was a survey more than 15 years ago by Biggs et al. in 1994).

This study has shown that FD initiatives can be an asset in recruiting and retaining teachers as they value the development opportunities. Competence and career progression were two of the four key outcomes reported by educators in the final theories (Table 9.9) and institutions need to note these outcomes. First is competence, it is important for institutions to have and retain competent faculty because at the heart of the university are its faculty members, the people who devote their lives to the teaching and service mission of the institution. Furthermore, the GMC (2012) document on the recognition and approval of trainers in the UK has now brought a mandatory component to FD provision in UK medical schools which

will necessitate universities having competent educators delivering teaching that enables students to achieve learning outcomes. Similarly, on the global stage, the WHO in its (2013) guideline on 'transforming and scaling up health professionals' education and training' recommended that health professionals' education and training institutions should design and implement programmes for faculty / teaching staff to promote teaching competency. It also emphasised the importance for governments, funders and accrediting bodies to support the implementation of higher education policies for mandatory FD programmes.

Second is the issue of career progression. This study has shown that educators highly valued their career progression following attendance at FD initiatives. This is similar to the findings of Steinert et al's (2009) study on faculty members' participation in FD in which the study participants believed that FD referred to their *own development* (personal and career development) as faculty members, and not merely the enhancement of specific competencies related to teaching. Interestingly, my literature review did not report many FD programmes focused on career development despite the fact that faculty members welcome the opportunity to identify career goals, form academic identities and acquire skills to further their career path (Wingard et al, 2004; Pololi et al, 2002; Miedzinski et al, 2001). Given that faculty members are the medical school's most important resource, it would seem that an investment in career development (both as a content area and as a strategy) through FD represents a critical step forward to enhance recruitment, promote retention, and create an environment that will enrich the academic role.

### 10.1.2 FD developers

What about FD developers, how does this realist study help them design their programmes? The first point to emphasise is that there is no single blueprint for FD development. Educators will consider a programme (or not), attend it (or not), learn lessons (or not), retain the lessons (or not), apply the lessons (or not) and each one of their decisions will be internally complex and depend on the educator's circumstances. FD developers have to understand that the real engine for change in a complex programme such as FD is the process of participants making constrained choices amongst the range of opportunities provided by the FD. This is why I recommend that FD developers use the realist *CMO* model (rather than focus only on process) to develop their programmes so that the desired outcomes are achieved by activating the right mechanisms in the appropriate contexts. They need to give contextual factors and causal mechanisms their proper places in the design process as shown in the final four theories. The four outcomes of confidence, competence, credibility and career progression are equally important to FD developers, as they should want their programmes to be successful in all domains. However in my opinion, if FD coordinators have to choose one theory then it should be Theory 1 (engagement), as without engagement there is no deep learning and consequently the other outcomes are not realised. While other authors have reached the same conclusion on engagement (Carpenter, 2010; Hargreaves, 2006), it is through the realist model that FD developers will understand the importance of incorporating the contextual factors (participatory, interactive, reflective practice) that need to be present for engagement to be activated.

The second point to consider is, if the goal of FD as articulated by the FD co-ordinators is to drive the standards of teaching and produce talented, professional educators to teach medical students (Box 8.1), then FD should be designed to reach that quality objective. The end result will be excellent delivery of patient care. Furthermore, programmes should be submitted to an accreditation process to ensure that quality is maintained and continuously enhanced but retain the flexibility to adapt programme to specific needs. For all these to happen, FD needs to become more tailored, contextualised and designed in collaboration with educators (and possibly by extension patients) as discussed in section 8.4.

### 10.1.3 Educators

Finally for the educators, what does this study add? First, it helps to address the issue of educators' attitude and misconceptions as a barrier to FD. As I discussed in the literature review (section 2.4.3), educators may over or under estimate their teaching ability, may not perceive the benefits of training or fail to recognise a link between teacher training and teaching excellence. Therefore, it can be argued that this study clearly provides educators with evidence of the positive outcomes of FD in terms of improving confidence, competence, credibility and career progression.

Secondly, the study provides educators with a different model to help them decide what works for them as they are often urged to use particular programmes or strategies in their practice. An example is the SPICES (student-centred, problem-based, integrated or interprofessional, community-based, elective driven and systematic) model for teaching which educators are still recommended to consider when developing courses (Harden et al, 1984). A closer look at the studies that

supported this strategy were inconclusive since the original research focused more on identifying outcomes and ignored the effect of context on those outcomes. Hence it would seem difficult to say with certainty that the SPICES strategy resulted in the specified outcomes. It would have been more useful for educators, if the researchers had been able to specify what aspects of the programme worked in what circumstances. Therefore, in my opinion the realist approach empowers educators to ask the question how, why and in what circumstances strategies would work.

Thirdly, educators could also use this realist study to evaluate aspects of their own everyday teaching. For example, an educator might want to examine how collaborative learning had worked for teaching his or her students. He or she could identify possible outcomes such as mastery of the subject content, development of communication and team working skills. The educator could then explore with the stakeholders (the university, the students and other teachers) which contexts (for example, heterogeneity or diversity of the group, time availability for tasks, interpersonal relationships and nature of the identified problem) facilitated successful outcomes and which hindered the mechanisms of the programme. In this way, the educator could then explain how different features of collaborative learning worked in various contexts.

I have summarised my recommendations for FD in Box 10.1

**Box 10.1: Recommendations for Faculty Development**

1. **Standards / Values:** Faculty development activities need to be certified or accredited for standardisation and professionalization of teaching. A baseline standard is needed for medical educators which have to be explicit and contractual with annual educational appraisal and annual faculty development refreshers. Faculty development needs to be valued at the institutional level with provision of excellent webpages, adequate resources and funding. Faculty development in teaching need to be valued as scholarship.

2. **Design:** Need for clarity of the theoretical approach used in faculty development programme design. There is a need for a more collaborative design model between faculty development programme developers and educators. Future faculty development opportunities have to be tailored and targeted to educators' needs as well as the provision of advanced programs (contextual or individualised) to achieve the desired outcomes (confidence, competence, credibility and career progression).

3. **Delivery:** This has to take into account the connections between contexts, mechanisms and outcomes in the four theories of FD (engagement, motivation, positive perception, and professionalization). Using teaching / learning strategies that motivate and engage learners to produce positive perceptions of the faculty development activity is important. Furthermore, faculty development coordinators and facilitators need to focus on the contexts (participatory approach, reflective practice, relevance and access to faculty development) that are vital to the activation of mechanisms in producing the desired outcomes.

4. **Evaluation:** The realist evaluation model provides a sound framework for the evaluation of faculty development programmes. There has to be an understanding of the connections between contexts, mechanisms and outcomes in the programme to provide the evidence of programme effectiveness for all stakeholders.

## 10.2 Recommendations for realist researchers in medical education

In the previous chapter (section 9.4), I discussed my modifications of the realist model for my study on FD in medical schools and shown that the model can be adapted for use in educational research. As previously alluded to, this realist model is more suitable for evaluating discrete programmes rather than policies. However, it can still be used to produce evidence on various aspects of professional practice, not just specific programmes. Examples could include the use of realist evaluation to assess the efficiency of a multi-agency meeting, or the usefulness of a professional

report. The principles of the evaluation would be the same: the professional would need to identify the outcomes and then use the stakeholders to explain what worked and in what circumstances. Matthews (2003) supported using realist methods as a way of producing the evidence base for professional practice and argued that realist evaluation, which allows the practitioner to explore the complexity of a phenomenon, is the most appropriate way to ensure that professional practice is of value to the relevant stakeholders. Therefore, an important contribution of this research inquiry is guidance for future medical education researchers on how the model can be used (summarised in Box 10.2). In essence, by adapting the realist evaluation model for use in medical education, my study has pioneered a method, which should have value for future researchers who wish to evaluate medical education programmes.

In the section after Box 10.2, I will describe my developmental journey as a realist researcher followed by my final conclusions on FD.

**Box 10.2: Recommendations for realist researchers in medical education**

1. **Purpose:** Determine clearly if the underlying purpose of the inquiry is to identify contexts that facilitate or inhibit causal mechanisms to produce various outcomes.

2. **Design:** Decide on the model to use. The realist evaluation model provides a sound structure for the evaluation of programmes in education research, particularly when the researcher is also a practitioner. For evidence-based policy the theories-of-change model is more appropriate.

3. **Hypotheses:** Review the literature, existing data, previous research when possible for evidence of possible contexts, mechanisms and outcomes i.e. 'good explanations' to underpin the construction of the hypotheses of the inquiry. Focus the area of study by limiting the number of hypotheses; only those that are vital to the functioning of the programme should be included.

4. **Data collection:** For determining methods of data collection, first clarify the information that is needed and then ask 'who can give me this information?' Finally ask, 'how can I collect this information?'

5. **Stakeholders:** Ensure that the group of stakeholders is sufficiently diverse to have knowledge of all aspects of the hypotheses. When discussing the hypotheses with the stakeholders, do this in a holistic way following the principles of theorising the interview.

6. **Data analysis:** This is where further modification becomes important. There is no predetermined realist analytic method. Summarising and displaying the data under each hypothesis is a good analytic choice. However, irrespective of the analytic approach, researchers need to use their understanding and insight to make the final deductions.

7. **Theory:** The final theory or theories should clearly show the *CMO* configuration i.e. the connections between contexts, mechanisms and outcomes in the programme evaluated.

## 10.3 My developmental journey as a realist researcher

In this section, I discuss my own journey as researcher who started out with some ideas about FD, albeit at the postgraduate medical education level, and my developmental journey through the various stages of the research. As discussed in section 1.2, I started my PhD journey as a clinician interested in medical education, lucky enough to have participated in the postgraduate training the trainers course in the West Midlands deanery, UK. Like most clinicians, I was trained in the quantitative paradigm and my initial foray into qualitative educational study was in outcome

based education (OBE). Therefore my initial research question was ‘whether FD is effective and whether it leads to good educational outcomes’. However, one year into the research study with wider reading on FD, on various learning theories, and guidance / feedback from my supervisors, I became a bit uncertain of my initial question. My supervisors encouraged me to dig deeper and explore the question further. My previous experience of FD gave me a positive bias that FD might be effective but the question became how and why? Hence, I started looking at a different aspect of my question.

My next challenge was finding a theory that would underpin my research study. I initially considered the constructivist epistemology with knowledge being constructed by the interaction between subject and object and meanings being constructed by human beings as they engage with the world they are interpreting. This epistemology with an interpretive framework (exploration of meaning) seemed to suit my purpose as I was interested in educators’ views on FD but I wasn’t convinced it would answer my question on how FD is effective. I had further dilemma as I debated the possibility of two theories: a macro design theory on which the research is based and a micro analytic theory, but in my view one theory for both aspects would work better. So was there an integrated theory that I could adopt? I explored transformation theory but, from my previous experience in postgraduate FD training, I was aware that most participants did not undergo an epochal transformative change; rather it was more of a modification of their teaching and delivery style. Next, I considered complexity theory as I viewed FD as operating in a complex system. Complexity theory concerns the behaviour of complex systems and processes with discrete elements interacting and patterns of order or behaviour emerging (Zimmerman et al, 1998;

Stacey, 2010). I also explored situated learning theory (Lave & Wenger, 1991) since it has more relation to context, but I wasn't convinced that any of the above theories would satisfy my understanding of why and how FD is effective.

It wasn't until 18 months after starting the PhD, and further discussion with one of my supervisors (JT) during a trip to Australia, that she encouraged me to consider realist evaluation and this crystallised my thoughts. It answered my query on the theory to use as it addressed both the process and outcome aspects of my question. The challenge was in designing it for medical education as it hasn't been widely used in this area, hence the modifications I described earlier. On reflection, even though at the time it didn't seem like it, this process of considering various theoretical approaches gave me a richer understanding of the qualitative research field bearing in mind my quantitative background.

The next stage in the research journey was grappling with the mixed methodology of data collection: webpage data, observation, and interviews with their individual nuances and challenges. The biggest challenge though was to come in the analysis stage as fitting the voluminous amount of data into the CMO configuration was challenging and time consuming. The various interpretations of contexts and mechanisms from an educational perspective weren't as clear cut compared with the physical world. However, with the help and support of my supervisors, I was able to arrive at four key theories which I believe are important for all FD stakeholders: the educators, the FD developers, the universities and policy makers.

In summary, my faculty development PhD journey has been a very positive and valuable experience. It has contributed immensely to my personal growth as a researcher with a lot more confidence in and appreciation of the qualitative paradigm.

## 10.4 Concluding Message

Medical educators need to be prepared for complex and demanding roles including teaching, scholarship, leadership, career development and FD initiatives play a vital role in this. Therefore, this innovative realist exploration of FD is particularly important in view of the major investment in time and resources for FD. I have shown FD to be effective once the contexts that activate the mechanisms leading to particular outcomes are understood. Moreover, realist evaluation of FD provides practitioners, developers and policy makers with an effective tool to evaluate FD in various circumstances and to develop evidence-based practice. Finally, I will modestly claim that my study has contributed to the development of medical education research and practice.

# REFERENCES


AAMC (2007) *Advancing Educators and Education: Defining the Components and Evidence of Educational Scholarship*. Washington, DC: Association of American Medical Colleges.

Alexander, J. E. and Tate, M. A. (1999) *Web Wisdom: How to Evaluate and Create Informational Quality on the Web*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Alexander, J. E. and Tate, M. A. (2001) *How to Recognise an Informational Web Page*. [online] Available from: www.widener.edu/about/campus_resources/wolfgram_library/evaluate (Accessed 20th May 2011).

Alvesson, M. and Sköldberg, K. (2009) *Reflexive Methodology: New Vistas for Qualitative Research*. Second. London: SAGE.

Amin, Z., Hoon Eng, K., Gwee, M. and Chay Hoon, T. (2006) Addresing the needs and priorities of medical teachers through a collaborative intensive faculty development programme. *Medical Teacher*, 28 (1): 85-88.

Atkinson, J. W. (1966) Motivational determinants of risk-taking behavior. In: Atkinson, J. W. and Feather, N. T. eds. *A theory of achievement motivation*. New York: John Wiley & Sons: 11 - 29.

Bakken, L. L. (2002) An evaluation plan to assess the process and outcomes of a learner-centered training program for clinical research. *Medical Teacher*, 24 (2): 162-168.

Bandura, A. (1997) *Self-efficacy: The exercise of control*. New York: Freeman.

Barlett, P. and Rappaport, A. (2009) Long term impact of faculty development programs. The experience of Teli and Piedmont. *College Teaching*, 57 (2): 73-82.

Baroffio, A., Kayser, B., Vermulen, B., Jacquet, J. and Vu, N. V. (1999) Improvement of tutorial skills: an effect of workshops or experience? *Academic Medicine*, 74 (10 Suppl.): S75-S77.

Baumeister, R. F. and Vohs, K. D. ed. (2004) *Handbook of Self-Regulation: Research, Theory, and Applications*. New York: Guilford.

Bazeley, P. (2007) *Qualitative Data Analysis with NViVo*. London: Sage.

Beattie, D. (2000) Expanding the View of Scholarship: Introduction. *Academic Medicine*, 75 (9): 871-876.

Becker, H. S. (1970) *Sociological work: Method and substance*. Chicago: Aldine.

Bennett, N., Crawford, M. and Riches, C. R. ed. (1992) *Managing change in education: Individual and organisational perspectives*. Milton Keynes: Open University Press.

Bernstein, D. A. (2013) *Essentials of Psychology*. 6th edition. Belmont, CA: Wadsworth Cengage Learning.

Bhaskar, R. (1975) *A realist theory of science*. Leeds: Leeds Book.

Bhaskar, R. (1989) *The Possibility of Naturalism*. Hemel Hempstead: Harvester Wheatsheaf.

Bhaskar, R. (2002) *From Science to Emancipation: Alienation and the Actuality of Enlightenment*. Thousand Oaks, CA: Sage.

Biggs, J. S. G., Agger, S. K., Dent, T. H. S., Allery, L. A. and Coles, C. (1994) Training for medical teachers: a UK survey. *Medical Education*, 28 (2): 99-106.

Blackwell, A., Bowes, L., Harvey, L., Hesketh, A. J. and Knight, P. T. (2001) Transforming Work Experience in Higher Education. *British Educational Research Journal*, 27 (3): 269-285.

Blaikie, N. (2000) *Designing Social Research: The Logic of Anticipation*. Cambridge: Polity Press.

Bland, C. J. and Froberg, D. G. (1982) A sytematic appraoch to faculty development for family practice faculty. *The Journal of Family Practice*, 14 (3): 537-543.

Bland, C. J. and Schmitz, C. (1986) Characteristics of the successful researcher and implications for faculty development. *Journal of Medical Education*, 1986 (61): 22-31.

Bland, C. J., Schmitz, C., Stritter, F., Aluise, J. and Henry, R. (1988) Project to identify essential faculty skills and develop model curricula for faculty development programs. *Journal of Medical Education*, 63 (6): 467-469.

Bland, C. J. and Schmitz, C. C. (1988) An overview of research on faculty and institutional vitality. In: Schuster, J. and Wheeler, D. eds. *Enhancing Faculty Careers: Strategies for Development and Renewal*. San Francisco, CA: Jossey-Bass: 41-61.

Bland, C. J., Schmitz, C. C., Stritter, F. T., Henry, R. C. and Alieve, J. J. (1990) *Successful faculty in academic medicine: essential skills and how to acquire them*. New York: Springer.

Bligh, D. (2000) *What's the use of lectures?* 5th edition. San Francisco, CA: Jossey-Bass.

Bligh, J. (2005) Faculty development. *Medical Education*, 39 (2): 120-121.

Boelen, C. (1999) Adapting health care institutions and medical schools to societies' needs. *Academic Medicine*, 74 (8): S11-20.

Boerboom, T. B., Dolmans, D. H., Muijtjens, A. M., Jaarsma, A. D., Van Beukelen, P. and Scherpbier, J. J. (2009) Does a faculty development programme improve teachers' perceived competence in different teacher roles? *Medical Teacher*, 31 (11): 1030-1031.

Bogdan, R. C. and Biklen, S. K. (2007) *Qualitative Research for Education: An Introduction to Theories and Methods*. 5th Edition. Boston, MA: Pearson Education.

Bohuijs, P. A. J. (1998) The Teacher and Self-Directed Learners. In: Jolly, B. and Rees, L. eds. *Medical Education in the Millennium*. Oxford: Oxford University Press: 192-198.

Boucher, B., Chyka, P., Fitzgerald, W., Hak, L., Miller, D., Parker, r. and Phelps, S. (2006) A comprehensive approach to faculty development. *American Journal of Pharmaceutical Education*, 70 (2): 1-6.

Boud, D. (2001) Introduction: Making the Move to Peer Learning. In: Boud, D., Cohen, R. and Sampson, J. eds. *Peer Learning in Higher Education: Learning From & With Each Other*. London: Kogan Page Ltd: 1-17.

Boud, D., Keogh, R. and Walker, D. (1985) *Reflection: turning experience into learning*. London: Kogan Page.

Boyd, R. D. and Myers, J. G. (1988) Transformative Education. *International Journal of Lifelong Education*, 17 (4): 261-284.

Boyer, E. (1990) *Scholarship Reconsidered: Priorities of the Professoriate*. New York: Jossey-Bass.

Branch, W., Frankel, R., Gracey, C., Haidet, P., Weissmann, P., Cantey, P., Mitchell, G. and Inui, T. (2009) A Good Clinician and a Caring Person: Longitudinal Faculty

Development and the Enhancement of the Human Dimensions of Care. *Academic Medicine*, 84 (1): 117-126.

Branch, W., Kroenke, K. and Levinson, W. (1997) The clinician-educator present and future roles. *Journal of General Internal Medicine*, 12 (Suppl 2): S1-S4.

Bransford, J., D, Brown, A., L and Cocking, R., R (2000) *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.

Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3 (2): 77-101.

Brissette, A. and Howes, D. (2010) Motivation In Medical Education: A Systematic Review *WebmedCentral Medical Education*, 1 (12): WMC001261.

Britten, N. (1995) Qualitative Research: Qualitative interviews in medical research. *BMJ*, 311 (6999): 251-253.

Bruer, J. T. (1993) *Schools of Thought: A Science of Learning in the Classroom*. Cambridge, MA: MIT Press.

Bryman, A. (2006) Paradigm peace and the implications for quality. *International Journal of Social Research Methodology*, 9 (2): 111-126.

Bryman, A. (2012) *Social research methods*. Oxford: Oxford University Press.

Bunniss, S. and Kelly, D. R. (2010) Research paradigms in medical education research. *Medical Education*, 44 (4): 358-366.

Busari, J. O., Scherpbier, A. J. J. A., van der Vleuten, C. P. M. and Essed, G. G. M. (2006) A 2-day teacher-training programme for medical residents: investigating the impact on teaching performance. *Advances in Health Sciences Education*, 11 (2): 133-144.

Byng, R., Norman, I. and Redfern, S. (2005) Using Realistic Evaluation to Evaluate a Practice-level Intervention to Improve Primary Healthcare for Patients with Long-term Mental Illness. *Evaluation*, 11 (1): 69-93.

Campbell, D. T. (1988) *Methodology and epistemology for social science: Selected papers*. Chicago: University of Chicago Press.

Cantillon, P. (2003) Teaching large groups. *BMJ*, 326 (7386): 437-440.

Carpenter, B. (2010) *A Vision for the 21st Century Special School*. London:

Carter, B. and New, C. ed. (2004) *Making Realism Work*. New York: Routledge.

Cartwright, N. (1999) *The dappled world: A study of the boundaries of science*. Cambridge, England: Cambridge University Press.

Centra, J. (1973) Effectiveness of student feedback in modifying college instruction. *Journal of Educational Psychology*, 65 (3): 395-401.

Centra, J. (1976) *Faculty development practices in U.S. colleges and universities.* Princeton, N.J: Educational Testing Service PE-76-30.

Chambers, R. (2013) PRA, PLA and Pluralism: Practice and Theory. In: Reason, P. and Bradbury, H. eds. *The Sage Handbook of Action Research: Participative Inquiry and Practice*. London: Sage: 297-318.

Chapple, A. (1999) The use of telephone interviewing for qualitative research. *Nurse Researcher*, 6; 85-93.

Cheetham, G. and Chivers, G. (1996) Towards a holistic model of professional competence. *Journal of European Industrial Training*, 20 (5): 20-30.

Cheetham, G. and Chivers, G. (1998) The reflective (and competent) practitioner. *Journal of European Industrial Training*, 22 (7): 267-276.

Clark, A. (2011) *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.

Clark, A. M., Hartling, L., Vandermeer, B. and McAlister, F. A. (2005) Meta-Analysis: Secondary Prevention Programs for Patients with Coronary Artery Disease. *Annals of Internal Medicine*, 143 (9): 659 - 672.

Clark, J., Houston, T., Kolodner, K., Branch, W., Levine, R. and Kern, D. (2004) National Survey of Faculty Development in Departments of Medicine of U.S. Teaching Hospitals. *Journal of General Internal Medicine*, 19 (3): 205-214.

Clark, S. M. and Lewis, D. R. (1985) Implications for institutional response. In: Clark, S. M. and Lewis, D. R. eds. *Faculty vitality and institutional productivity*. New York: Teachers College Press: 55-72.

Clarke, D. D. (2004) Structured judgement methods - the best of both worlds? In: Todd, Z., Nerlich, B., Mckeown, S. and Clarke, D. D. eds. *Mixing Methods in Psychology*. Hove, East Sussex: Psychology Press: 79-98.

Cochran, L. H. (1989) Administrative commitment to leaching: Practical, research-based strategies to strengthen college leaching effectiveness. *ERIC*, 318 (334): 9.

Coffey, A. and Atkinson, P. (1996) *Making Sense of Qualitative Data*. Thousand Oaks, CA: Sage.

Cohen, D. J. and Crabtree, B. F. (2006) *Qualitative research guidelines project*. [online] Available from: http://www.qualres.org/HomeEval-3664.html (Accessed June 2010).

Cohen, L., Manion, L. and Morrison, K. (2011) *Research Methods in Education*. 7th edition. London: Routledge.

Cohen, R., Murnaghan, L., Collins, J. and Pratt, D. (2005) An update on master's degrees in medical education. *Medical Teacher*, 27 (8): 686-692.

Colthart, I., Bagnall, G., Evans, A., Allbutt, H., Haig, A., Illing, J. and McKinstry, B. (2008) The effectiveness of self-assessment on the identification of learner needs, learner activity, and impact on clinical practice: BEME Guide no. 10. *Medical Teacher*, 30 (2): 124-145.

Cooke, M., Irby, D. and Debas, H. (2003) The UCSF Academy of Medical Educators. *Academic Medicine*, 78 (7): 666-672.

Coon, D. and Mitterer, J. O. (2009) *Introduction to Psychology: Gateways to Mind and Behavior.* Belmont, CA: Wadsworth Cengage Learning.

Cornford, J. and Pollock, N. (2002) *Putting the University Online: Information, Technology and Organisational Change*. Buckingham: Open University Press.

Coughlan, S. (2008) *"Faculty in league table expulsion"*. BBC News (25.07.2008).

Cox, A. (2007) The Power and Vulnerability of the 'New Professional': Web Management in UK Universities. *Program: electronic library and information systems*, 41 (2): 148-169.

Cox, A. and Emmott, S. (2007) A survey of UK University Web Management: Staffing Systems and Issues. *Campus-Wide Information Systems*, 24 (5): 308-330.

Creswell, J. W. (2007) *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.

Crotty, M. (1998) *The Foundations of Social Research*. Thousand Oaks, CA: Sage.

Cruess, S. R., Johnston, S. and Cruess, R. L. (2004) Profession: A Working Definition for Medical Educators. *Teaching and Learning in Medicine*, 16 (1): 74-76.

CUG (2012) *Complete University Guide*. [online] Available from: www.thecompleteuniversityguide.co.uk/league-tables/rankings (Accessed 18.05.2011).

D'Eon, M., Overgaard, V. and Harding, S. R. (2000) Teaching as social practice: Implications for faculty development. *Advances in Health Sciences Education:Theory and Practice*, 5 (2): 151-162.

Danermark, B., Ekström, M., Jakobsen, L. and Karlsson, J. C. (2002) *Explaining Society: Critical realism in the social sciences*. London: Routledge.

Davies, P. (2007) The radical accuracy of the subjective viewpoint. *British Journal of General Practice*, 57 (543): 848-849.

de Montaigne, M. (2004) *The Complete Essays*. London: Penguin Classics (originally published 1588).

Dearing, R. (1997) Higher Education in the Learning Society: National Committee of Inquiry into Higher Education.

Delamont, S. (2002) *Fieldwork in Education Settings: Methods, Pitfalls and Perspectives*. 2nd edition. New York: Routledge.

Dennick, R. (2003) Long-term retention of teaching skills after attending the Teaching Improvement Project: a longitudinal, self-evaluation study. *Medical Teacher*, 25 (3): 314-318.

Dennick, R. G. (1998) Teaching medical educators to teach: the structure and participant evaluation of the Teaching Improvement Project. *Medical Teacher*, 20 (6): 598-601.

Dewey, J. (1933) *How we think: A restatement of the relation of reflective thinking to the educative process*. Boston, Massachusetts: Heath.

Dexter, L. A. (1970) *Elite and specialised interviewing*. Evanston, IL: Northwestern University Press.

DiCicco-Bloom, B. and Crabtree, B. F. (2006) The qualitative research interview. *Medical Education*, 40 (4): 314-321.

Donald, J. G. (1977) Research report. The search for teaching competencies in higher education. *Faculty Development and Evaluation Newspaper, 3(2),* Date: 11-14.

Dorene, F. B. and Boyd, F. R. (2012) Faculty Development as Transformation: Lessons Learned From a Process-Oriented Program. *Teaching and Learning in Medicine: An International Journal*, 24 (3): 242-247.

Drescher, U., Warren, F. and Norton, K. (2004) Towards evidence based practice in medical training: making evaluations more meaningful. *Medical Education*, 38 (12): 1288-1294.

Dweck, C. S. and Molden, D. C. (2005) Self-theories: Their impact on competence motivation and acquisition. In: Elliot, A. J. and Dweck, C. S. eds. *Handbook of competence and motivation* New York: Guilford Press: 120-140.

Eble, K. E. and McKeachie, W. J. (1985) *Improving undergraduate education through faculty development.* San Francisco, CA: Jossey-Bass.

Eccles, J. S. and Wigfield, A. (2002) Motivational beliefs, values and goals. *Annual Review of Psychology*, 53 (1): 109-132.

Eduserv (2009) *Investigation into the management of website content in higher education institutions*. Social Issues Research Centre http://www.eduserv.org.uk/,

Eitel, F., Kanz, K. G. and Tesche, A. (2000) Training and certification of teachers and trainers: The professionalisation of medical education. *Medical Teacher*, 22 (5): 517-526.

Elder-Vass, D. (2010) *The Causal Power of Social Structure*. Cambridge: Cambridge University Press.

Elton, L. (1991) Teaching Excellence and Quality Assurance. *Zeitschrift für Hochschuldidaktik*, 1 (2): 102-115.

Emerson, R. M., Fretz, R. I. and Shaw, L. L. (2011) *Writing ethnographic fieldnotes*. 2nd edition. Chicago, IL: University of Chicago Press.

Entwistle, N. J., McCune, V. and Walker, P. (2001) Conceptions, Styles and Approaches Within Higher Education: Analytic Abstractions and Everyday Experience. In: Sternberg, R. and Zhang, L.-F. eds. *Perspectives on Thinking, Learning and Cognitive Styles*. Mahwah, N.J: Lawrence Erlbaum Associates: 85-115.

Entwistle, N. J. and Ramsden, P. (1983) *Understanding Student Learning.* London: Croom Helm.

Epstein, R. M. (2008) Reflection, perception and the acquisition of wisdom. *Medical Education*, 42 (11): 1048-1050.

Erickson, F. (1986) Qualitative methods. In: Wittrock, M. eds. *Handbook of research on teaching*. New York: Macmillan: 119-161.

Eva, K. W. and Regehr, G. (2005) Self-Assessment in the Health Professions: A Reformulation and Research Agenda. *Academic Medicine*, 80 (10 Suppl): S46 - 54.

Eva, K. W. and Regehr, G. (2011) Exploring the divergence between self-assessment and self-monitoring. *Advances in Health Sciences Education*, 16 (3): 311-329.

Falchikov, N. (1995) Improving feedback to and from students. In: Knight, P. eds. *Assessment for Learning in Higher Education (Staff & Educational Development)*. Oxon: Routledge: 157-166.

Fidler, D., Khakoo, R. and Miller, L. A. (2007) Teaching Scholars Programs: Faculty Development for Educators in the Health Professions. *Academic Psychiatry*, 31 (6): 472-478.

Finucane, P., Allery, L. A. and Hayes, T. M. (1994) Attitudes to teaching among teachers at a British medical school. *Medical Education*, 28 (3): 213-219.

Fishbein, M. and Ajzen, I. (1975) *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Francis, J. B. (1975) How do we get there from here? Program design for faculty development. *Journal of Higher Education*, 46 (6): 719-732.

Fredericks, J. A., Blumenfeld, P. and Paris, A. H. (2004) School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74 (1): 59-109.

Freeman, J., Loewe, R. and Benson, J. (1998) Training Family Medicine Faculty to Teach in Underserved Settings. *Family Medicine*, 30 (3): 168-172.

Freeth, D., Hammick, M., Koppel, I., Reeves, S. and Barr, H. (2002) *A Critical Review of Evaluations of Interprofessional Education*. [online] Available from: London, Higher Education Academy Learning and Teaching Support Network for Health Sciences and Practice,

http://www.health.ltsn.ac.uk/publications/occasionalpaper/occasionalpaper02.pdf (Accessed 21.03.2012).

Frey, J. and Fontana, A. (1991) The group interview in social research. *The Social Science Journal*, 28 (2): 175-187.

Frohna, A., Hamstra, S., Mullan, P. and Gruppen, L. (2006) Teaching Medical Education Principles and Methods to Faculty Using an Active Learning Approach: The University of Michigan Medical Education Scholars Program. *Academic Medicine*, 81 (11): 975-978.

Fry, H. (2006) Professional development for teaching in Higher Education: a brief account of the evolution of accredited programmes in the UK. *Zeitschrift fu¨r Hochschulentwicklung*, 1 (2): 1-13.

Gaff, J. G. (1975) *Toward faculty renewal: Advances in faculty, instructional, and organizational development.* San Francisco, CA: Jossey-Bass.

Gaff, J. G. and Simpson, R. (1994) Faculty development in the United States. Innovative Higher Education. *Innovations in Higher Education*, 18 (3): 167-176.

General Medical Council (2002) *Tomorrow's Doctors, Recommendations on Undergraduate Medical Education*. London: General Medical Council.

General Medical Council (2010) *GMC Education Strategy 2011-2013*. London: General Medical Council.

Gill, M. and Turbin, V. (1999) Evaluating "Realistic Evaluation": Evidence from a Study of CCTV In: Painter, K. and Tilley, N. eds. *Surveillance of Public Space: CCTV, Street Lighting and Crime Prevention*. New York: Criminal Justice Press: 179 -199.

Glanville, J. L. and Wildhagen, T. (2007) The measurement of school engagement: Assessing dimensionality and measurement invariance across race and ethnicity. *Educational and Psychological Measurement*, 67: 1019 -1041.

GMC (2003) *Tomorrow's Doctors. Report of the Education Committee of the GMC*. London: General Medical Council.

GMC (2006) *Good Medical Practice*. London: General Medical Council.

GMC (2009) *Tomorrow's Doctors*. London: General Medical Council.

GMC (2012) *Recognising and Approving trainers: The implementation plan*. London: General Medical Council.

Godfrey, J., Dennick, R. and Welsh, C. (2004) Training the trainers: do teaching courses develop teaching skills? *Medical Education*, 38 (8): 844-847.

Goldszmidt, M., Zibrowski, E. and Weston, W. W. (2008) Education scholarship: it's not just a question of 'degree'. *Medical Teacher*, 30 (1): 34-39.

Gonczi, A. (1999) Competency-based learning: a dubious past - an assured future? In: Boud, D. and Garrick, J. eds. *Understanding Learning at Work*. London: Routledge: 180 - 196.

Good, T. L. and Brophy, J. E. (1990) *Educational psychology: A realistic approach*. (4th ed.). White Plains, NY: Longman.

Gosling, D. (2002) *Models of peer observation of teaching*. [online] Available from: www.heaacademy.co.uk (Accessed 25 June 2010).

Grabov, V. (1997) The Many Facets of Transformative Learning Theory and Practice. In: Cranton, P. eds. *Transformative Learning in Action: Insights from Practice. New Directions for Adult and Continuing Education. No 74*. San Francisco, CA: Jossey-Bass: 89-96.

Graneheim, U. H. and Lundman, B. (2004) Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, 24 (2): 105-112.

Greene, C., Lehn, B. and Goodyear, L. (2001) The merits of mixing methods in evaluation. *Evaluation*, 7 (1): 25-44.

Greene, J. (2007) *Mixed methods in social inquiry*. San Francisco, CA: Jossey-Bass.

Greenhalgh, T., Humphrey, C., Hughes, J., Macfarlane, F., Butler, C. and Pawson, R. (2009) How Do You Modernize a Health Service? A Realist Evaluation of Whole-Scale Transformation in London. *Milbank Quarterly*, 87 (2): 391-416.

Gregory, R. L. ed. (1987) *Oxford Companion to the Mind.* Oxford: Oxford University Press.

Gruppen, L., Frohna, A., Anderson, R. and Lowe, K. (2003) Faculty Development for Educational Leadership and Scholarship. *Academic Medicine*, 78 (2): 137-141.

Gruppen, L., Simpson, D., Searle, N., Robins, L., Irby, D. and Mullan, P. (2006) Educational fellowship programs: Common themes and overarching issues. *Academic Medicine*, 81 (11): 990-994.

Guskey, T. K. (2003) What makes professional development effective? *Phi Delta Kappan*, 84 748-750.

Hafferty, F. (2008) In search of a Lost Chord: Professionalism and medical education's hidden curriculum. In: Wear, D. and Bickel, J. eds. *Educating for professionalism: Creating a culture of humanism in medical education*. Iowa City, IA: University of Iowa Press: 11-34.

Hall, T. M. and Walck, C. (2007) Change Issues in Curriculum and Instruction / Credibility http://www.en.wikibooks.org/wiki/changeissuesincurriculumandinstruction/credibility

Hammersley-Fletcher, L. and Orsmond, P. (2004) Evaluating our peers: is peer observation a meaningful process? *Studies in Higher Education*, 29 (4): 489-503.

Hammersley, M. (1990) *Reading Ethnograpic research: a Critical Guide*. London: Longmans.

Hammersley, M. (1992) *What's wrong with ethnography? Methodological explorations*. London: Routledge.

Hammersley, M. and Atkinson, P. M. P. (2007) *Ethnography: principles in practice*. 3rd edition. New York: Routledge.

Harden, R. M. (1999) Stress, pressure and burnout in teachers: is the swan exhausted? *Medical Teacher*, 21 (3): 245-247.

Harden, R. M. (2002) Developments in outcome-based education. *Medical Teacher*, 24 (2): 117-120.

Harden, R. M., Crosby, J. R. and Davis, M. H. (1999) AMEE Guide No. 14: Outcome-based education: Part 1 - An introduction to outcome-based education. *Medical Teacher*, 21 (1): 7-14.

Harden, R. M. and Laidlaw, J., M (2012) *Essential skills for a medical teacher*. London: Churchill Livingstone.

Harden, R. M., Sowden, S. and Dunn, W. R. (1984) Educational strategies in curriculum development: the SPICES model. *Medical Education*, 18 (4): 284-297.

Hargreaves, D. (2006) Personalising Learning 6: the final gateway: school design and organisation.

Harris, D. L., Krause, K. C., Parish, D. C. and Smith, M. U. (2007) Academic competencies for medical faculty. *Family Medicine*, 39 (5): 343-350.

Harrison, C. and Hogg, W. (2003) Why do doctors attend traditional CME events if they don't change what they do in their surgeries? Evaluation of doctors' reasons for attending a traditional CME programme. *Medical Education*, 37 (10): 884-888.

Hartman, S. and Nelson, M. (1992) What we say and what we do: Self-reported teaching behavior versus performances in written simulations among medical school faculty. *Academic Medicine*, 67 (8): 522-527.

Haskins, W. (2000) Ethos and Pedagogical Communication: suggestions for enhancing credibility in the classroom. *Current Issues in Education*, 3 (4): 1-3.

Hativa, N., Barak, R. and Simhi, E. (2001) Exemplary university teachers. Knowledge and beliefs regarding effective teaching dimensions and strategies *The Journal of Higher Education*, 72 (6): 699-729.

Hattie, J. and Timperley, H. (2007) The Power of Feedback. *Review of Educational Research*, 77 (1): 81-112.

Hatton, N. and Smith, D. (1995) Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, 11 (1): 33-49.

HEA (2006) *Higher Education Academy*. [online] Available from: https://www.heacademy.ac.uk (Accessed 20.09.2010).

HEA (2011) *Higher Education Academy, Revised UK Professional Standards Framework* [online] Available from: https://www.heacademy.ac.uk (Accessed 12.12. 2013).

Healey, M. (2000) Developing the scholarship of teaching in Higher Education: a discipline-based approach. *Higher Education Research Development*, 19 (2): 169-189.

Herrmann, M., Lichte, T., Von Unger, H., Gulich, M., Waechtler, H., Donner-Banzhoff, N. and Wilm, S. (2007) Faculty development in general practice in Germany: experiences, evaluations, perspectives. *Medical Teacher*, 29 (2-3): 219-224.

Hesketh, E. A., Bagnall, G., Buckley, E. G., Friedman, M., Goodall, E., Harden, R. M., Laidlaw, J. M., Leighton-Beck, L., McKinley, P., Newton, R. and Oughton, R. (2001) A framework for developing excellence as a clinical educator. *Medical Education*, 35 (6): 555-564.

Hewson, M. (2000) A theory based faculty development program for clinician-educators. *Academic Medicine*, 75 (5): 498-501.

Hewson, M. and Copeland, L. (1999) Outcomes assessment of a faculty development program in medicine and pediatrics. *Academic Medicine*, 74 (10 Suppl): S68-S71.

Higgs, J. and McAllister, L. (2007a) Educating clinical educators: using a model of the experience of being a clinical educator. *Medical Teacher*, 29 (2): e51-e57.

Higgs, J. and McAllister, L. (2007b) Being a clinical educator. *Advances in Health Sciences Education*, 12 (2): 187-199.

Higgs, P., Jones, I. R. and Scrambler, G. (2004) Class as variable, class as generative mechanism. In: Carter, B. and New, C. eds. *Making Realism Work*. New York: Routledge: 91-110.

Hill, F. and Stephens, C. (2005) Building leadership capacity in medical education: developing the potential of course coordinators. *Medical Teacher*, 27 (2): 145-149.

Hitchcock, M. A., Stritter, F. T. and Bland, C. J. (1993) Faculty development in the health professions: conclusions and recommendations. *Medical Teacher*, 14 (4): 295-309.

Hollenberg, E., Reeves, S., Beduz, M., Jeffs, L., Kwan, D., Lee, J., Lowe, M., Merkley, J., Sinclair, L., Tassone, M. and Oandasan, I. (2009) Mainstreaming Interprofessional Education within Hospital Settings: Findings from a Multiple Case Study. *Journal of Research in Interprofessional Practice and Education*, 1 (1): 10-23.

Holton, E. F. (1996) The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7 (1): 5-21.

Honey, P. and Mumford, A. (1992) *The Manual of Learning Styles*. 3rd Edn. Maidenhead, London: Peter Honey Publications.

Horkheimer, M. (1972) *Critical Theory: Selected Essays*. New York: Herder and Herder.

Houlden, R. L. and Collier, C. P. (1999) Learning outcome objectives: A critical tool in learner-centered education. *journal of Continuing Education in the Health Professions*, 19 (4): 208-213.

Hume, D. (1978) *A treatise of human nature*. 2nd Edition. Oxford, England: Oxford University Press.

Institute for International Medical Education (2002) Global minimum essential requirements in medical education. *Medical Teacher*, 24 (2): 130-135.

Issenberg, S. B., Mcgaghie, W. C., Petrusa, E. R., Gordon, D. L. and Scalese, R. J. (2005) Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher*, 27 (1): 10-28.

Jonassen, D. H. (1991) Objectivism versus constructivism: do we need a new philosophical paradigm? *Educational Technology Research and Development*, 39 (3): 5-14.

Kennedy G. A (Translator) (1991) *The Rhetoric of Aristotle: A Theory of Civic Discourse*. Oxford: Oxford University Press.

Kennedy, T., Regehr, G., Rosenfeld, J., Roberts, W. and Lingard, L. (2004) Exploring the Gap Between Knowledge and Behavior: A Qualitative Study of Clinician Action Following an Educational Intervention. *Academic Medicine*, 79 (5): 386-393.

Killion, J. P. and Todnem, G. R. (1991) A process of personal theory building. *Educational Leadership*, 48 (6): 14-17.

Kirkpatrick, D. and Kirkpatrick, W. (2009) *Kirkpatrick, Then and Now: A Strong Foundation for the Future*. St Louis, MO: Kirkpatrick Partners.

Kirkpatrick, D. L. (1967) Evaluation of training. In: Craig, R. and Bittel, L. eds. *Training and Development Handbook*. New York, NY: McGraw-Hill: 87-112.

Kirkpatrick, D. L. (1994) *Evaluating Training Programs: The Four Levels*. San Francisco, CA: Berrett-Koehler Publishers.

Knight, A., Carrese, J. and Wright, S. (2007) Qualitative assessment of the long-term impact of a faculty development programme in teaching skills. *Medical Education*, 41 (6): 592-600.

Knight, P. T. and Trowler, P. R. (2000) Department-level cultures and the improvement of learning and teaching. *Studies in Higher Education*, 25 (1): 69-83.

Knowles, M. (1978) *The adult learner: A neglected species* 2nd Oxford, England: Gulf Publishing.

Knowles, M. S. (1980) *The Modern Practice of Adult Education: From Pedagogy to Andragogy.* Englewood Cliffs, NJ: Prentice Hall Reagents.

Knowles, M. S., Holton, E. and Swanson, R. A. (1998) *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. 5th ed. Houston, Texas: Gulf Publishing Company.

Kolb, D. A. (1984) *Experiential Learning: experience as the source of learning and development*. Upper Saddle River, NJ: Prentice-Hall.

Koppel, I., Barr, H., Reeves, S., Freeth, D. and Hammick, M. (2001) Establishing a systematic approach to evaluating the effectiveness of interprofessional education. *Issues in Interdisciplinary Care*, 3 (1): 41-49.

Koshy, V. (2005) *Action Research for Improving Practice: A Practical Guide*. London: Paul Chapman.

Kottkamp, R. B. (1990) Means for facilitating reflection. *Education and Urban Society*, 22 (2): 182-203.

Kough, K. (1997) *Communication in the classroom*. Prospect Heights, IL: Wavelane Press, Inc.

Krippendorff, K. (2004) *Content Analysis: An Introduction to its Methodology*. 2nd Edition. Thousand Oaks, CA: Sage.

Kua, E. H., Voon, F., Tan, C. H. and Goh, L. (2006) What makes an effective medical teacher? Perceptions of medical students. *Medical Teacher*, 28 (8): 738-741.

Kvale, S. and Brinkmann, S. (2009) *Interviews: Learning the Craft of Qualitative Research Interviewing*. Second edition. London: Sage.

Lakoff, G. (1987) *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Landis, R. B. (2000) *Studying Engineering: A Road Map to a Rewarding Career*. Second edition. Los Angeles, CA: Discovery Press.

Lang, F., Everett, K., McGowen, R. and Bennard, B. (2000) Faculty Development in Communication Skills Instruction: Insights from a Longitudinal Program with real-time feedback. *Academic Medicine*, 75 (12): 1222-1228.

Lave, J. and Wenger, E. (1991) *Situated Learning. Legitimate peripheral participation*. Cambridge: University of Cambridge Press.

Lawson, T. (1997) *Economics and Reality*. London: Routledge.

Layder, D. (1990) *The Realist Image in Social Science*. London: Macmillan.

Leinster, S. (2002) Medical education and the changing face of health care delivery. *Medical Teacher*, 24 (1): 13-15.

Lewis, K. G. (1996) Faculty development in the United States: A brief history. *International Journal for Academic Development*, 1 (2): 26-33.

Liben, S., Steinert, Y., Meterissian, S., Boillat, M. and McLeod, P. J. (2005) *Assessing faculty needs for faculty development: continuity or change?* Paper Presented at the Annual Meeting of the Canadian Association for Medical Education; Saskatoon, Saskatchewan.

Light, G. and Calkins, S. (2008) The experience of faculty development: patterns of variation in conceptions of teaching. *International Journal for Academic Development*, 13 (1): 27-40.

Little, D. (2010) *New contributions to the philosophy of history*. New York: Springer.

Litzelman, D. K., Stratos, G. A., Marriot, D. J., Lazaridis, E. N. and Skeff, K. M. (1998) Beneficial and harmful effects of augmented feedback on physicians' clinical-teaching performances. *Academic Medicine*, 73 (3): 324-332.

Lleras, C. (2005) Path analysis. *Encyclopedia of Social Measurement*, 3: 25-30.

MacDougall, J. and Drummond, M. J. (2005) The development of medical teachers: an enquiry into the learning histories of 10 experienced medical teachers. *Medical Education*, 39 (12): 1213-1220.

Manicas, P. T. (2009) Realist metatheory and qualitative methods. *Sociological Analysis*, 33 (1): 31-46.

Marchal, B., van Belle, S., van Olmen, J., Hoerée, T. and Kegels, G. (2012) Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18 (2): 192-212.

Martin, A. J. (2007) Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology*, 77 (2): 413-440.

Martin, A. J. (2008) Enhancing student motivation and engagement: The effects of a multidimensional intervention. *Contemporary Educational Psychology*, 33 (2): 239-269.

Marvel, M. K. (1991) Improving clinical teaching skills using the parallel process model. *Family Medicine*, 23 (4): 279-284.

Matthews, J. (2003) A framework for the creation of practitioner-based evidence. *Educational and Child Psychology*, 20 (4): 60-67.

Maxwell, J. A. (2012) *A Realist Approach for Qualitative Research*. London: Sage.

Maxwell, J. A. (2013) *Qualitative Research Design: An Interactive Approach*. 3rd edition. London: Sage.

McCracken, G. (1988) *The Long Interview*. Newbury Park, CA: Sage.

McCrosky, J. and Young, T. (1981) Ethos and credibility: The construct and its measurements after two decades. *Central States Speech Journal*, 32 (1): 24-34.

McGaghie, W. C., Bogdewic, S., Reid, A., Arndt, J. E., Stritter, F. T. and Frey, J. J. (1990) Outcomes of a faculty development fellowship in family medicine. *Family Medicine*, 22 (3): 196-200.

McIntyre, A. (2008) *Participatory Action Research*. Thousand Oaks, CA: Sage.

McLaughlin, S. (2005) Faculty development. *Academic Emergency Medicine*, 12 (4): 302e

McLean, M., Cilliers, F. and Van Wyk, J. (2008) Faculty development: Yesterday, today and tomorrow. *Medical Teacher*, 30 (6): 555-584.

McLeod, P., Steinert, Y., Meagher, T., Schuwirth, L., Tabatabai, D. and McLeod, A. H. (2006) The acquisition of tacit knowledge in medical education: learning by doing. *Medical Education*, 40 146-149.

McLeod, P. J. (1987) Faculty development practices in Canadian medical schools. *Canadian Medical Association Journal*, 136 (1): 709-712.

McLeod, P. J. and McLeod, A. H. (2004) Personal ViewIf formal CME is ineffective, why do physicians still participate? *Medical Teacher*, 26 (2): 184-186.

Mcleod, P. J. and Steinert, Y. (2010) The evolution of faculty development in Canada since the 1980s: Coming of age or time for a change? *Medical Teacher*, 32 (1): e31-e35.

McLeod, P. J., Steinert, Y., Nasmith, L. and Conochie, L. (1997) Faculty development in Canadian medical schools: a 10-year update. *Canadian Medical Association Journal*, 156 (10): 1419-1423.

McManus, I. C. (2003) Commentary: How to derive causes from correlations in educational studies. *British Medical Journal*, 326: 432.

McManus, I. C., Keeling, A. and Paice, E. (2004) Stress, burnout and doctors' attitudes to work are determined by personality and learning style: A twelve year longitudinal study of UK medical graduates. *BMC Medicine*, 2 (1): 29.

McNiff, J., Lomax, P. and Whitehead, J. (2003) *You and Your Action Research Project*. 2nd Edition. London: RoutledgeFalmer.

Medical Schools Council (2010a) *Medical schools in the UK, 2010*. [online] Available from: http://www.medschools.ac.uk (Accessed 15th July 2010).

Medical Schools Council (2010b) *Survey of Staffing Levels of Medical Clinical Academics in UK Medical Schools.* [online] Available from: http://www.medschools.ac.uk (Accessed 20th July 2010).

Menzel, H. (1978) Meaning: Who need it? In: Brenner, M., Marsh, P. and Brenner, M. eds. *The social contexts of methods*. New York, NY: St, Martin's Press: 140-171.

Merriam, S. B. and Caffarella, R. S. (1999) *Learning in Adulthood: A Comprehensive Guide*. 2nd. San Francisco: Jossey-Bass Publishers.

Merton, R. (1967) *On Theoretical Sociology. Five Essays, Old and New*. New York: The Free Press.

Mezirow, J. (1978) Perspective Transformation. *Adult Education*, 28: 100-110.

Mezirow, J. (2000) *Learning as Transformation: Critical Perspectives on a Theory in Progress.* San Francisco, CA: Jossey Bass.

Miedzinski, L. J., Davis, P., Al-Shurafa, H. and Morrison, J. C. (2001) A Canadian faculty of medicine and dentistry's survey of career development needs. *Medical Education*, 35 (9): 890-900.

Miles, M. B., Huberman, A. M. and Saldana, J. (2014) *Qualitative Data Analysis: A Methods Sourcebook*. 3rd Edition. Thousands Oak, CA: Sage.

Miliszewska, I. and Horwood, J. (2004) Engagement theory: A framework for supporting cultural differences in transnational education.

Miller, G. (1990) The assessment of clinical skills, competence and performance. *Academic Medicine*, 65 (9): S63 - S67.

Misch, D. A. (2002) Andragogy and medical education: Are medical students internally motivated to learn? *Advances in Health Sciences Education*, 7 (2): 153-160.

Mohanna, K., Wall, D. and Chambers, R. (2004) *Teaching made easy: a manual for health professionals.* 2nd Oxford: Radcliffe Medical Press.

Mohr, L. B. (1982) *Explaining organisational behaviour*. San Francisco: Jossey-Bass.

Moon, J. A. (1999) *Reflection in learning and professional development:Theory and practice*. London: Kogan Page.

Morrison, E. H., Rucker, L., Boker, J. R., Hollingshead, J., Hitchcock, M. A., Prislin, M. D. and Hubbell, F. A. (2003) A pilot randomised, controlled trial of a longitudinal residents-as-teachers curriculum. *Academic Medicine*, 78 (7): 722-729.

Moyles, J. (2002) Observation as a research tool. In: Coleman, M. and Briggs, A. J. eds. *Research Methods in Educational Leadership*. London: Paul Chapman: 172-191.

Muller, J. and Irby, D. (2006) Developing Educational Leaders: The Teaching Scholars Program at the University of California, San Francisco, School of Medicine. *Academic Medicine*, 81 (11): 959-964.

Nelson, W. (1983) Faculty who stay: Renewing our most important resources. In: Baldwin, R. and Blackburn, R. eds. *College faculty: Versatile human resources in a period of constraint*. San Francisco, CA: Jossey-Bass: 67- 83.

Newland, M., Newland, J., Steele, D., Lough, D. and McCurdy, F. (2003) Experience with a program of faculty development. *Medical Teacher*, 25 (2): 207-209.

Newswander, L. K. and Borrego, M. (2009) Engagement in two interdisciplinary graduate programs. *Higher Education*, 58 (4): 551-562.

Norton, L., Richardson, J. T. E., Hartley, J., Newstead, S. and Mayes, J. (2005) Teachers' beliefs and intentions concerning teaching in higher education. *Higher Education*, 50 (4): 537-571.

Notzer, N. and Abramovitz, R. (2008) Can brief workshop improve clinical instruction? *Medical Education*, 42 (2): 152-156.

Ogrinc, G. and Batalden, P. (2009) Realist Evaluation as a Framework for the Assessment of Teaching About the Improvement of Care. *Journal of Nursing Education*, 48 (12): 661-667.

Ogrinc, G., Headrick, L. and Boex, J. (1999) Understanding the value added to clinical care by educational activities. *Academic Medicine*, 74 (10): 1080-1086.

Orlander, J., Gupta, M., Fincke, G., Manning, E. and Hershman, W. (2000) Co-teaching: a faculty development strategy. *Medical Education*, 34 (4): 257-265.

Palmgreen, P. and Rayburn, J. D. (1985) An expectancy-value approach to media gratification. . In: Rosengren, K. E., Wenner, L. A. and Palmgreen, P. eds. *Media gratifications research: Current perspectives*. Beverly Hills, CA: Sage: 61-72.

Palys, T. (2008) Purposive sampling. In: Given, L. eds. *The SAGE encyclopedia of qualitative research methods*. Thousand Oaks: CA: Sage: 697-698.

Pandachuck, K., Harley, D. and Cook, D. (2004) Effectiveness of a Brief Workshop Designed to Improve Teaching Performance at the University of Alberta. *Academic Medicine*, 79 (8): 798-804.

Patton, M. (1997) *Utilization-Focused Evaluation: The New Century Text*. Thousand Oaks, CA: Sage.

Patton, M. Q. (1990) *Qualitative evaluation and research methods*. Newbury Park, London: Sage.

Patton, M. Q. (2002) *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

Patton, M. Q. (2011) *Developmental evaluation*. New York: Guilford Press.

Pawson, R. (1996) Theorising the Interview. *The British Journal of Sociology*, 47 (2): 295-314.

Pawson, R. (2002a) Does Megan's Law work? A theory-driven systematic review. *ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 8*, London: Queen Mary, University of London

Pawson, R. (2002b) Evidence-based Policy: In Search of a Method. *Evaluation*, 8 (2): 157-181.

Pawson, R. (2002c) Evidence-based Policy: The Promise of `Realist Synthesis'. *Evaluation*, 8 (3): 340-358.

Pawson, R. (2003a) Social care knowledge: seeing the wood for the trees. *ESRC UK Centre for Evidence-Based Policy and Practice: Working Paper 12*, London: Queen Mary, University of London

Pawson, R. (2003b) Nothing as Practical as a Good Theory. *Evaluation*, 9 (4): 471-490.

Pawson, R. (2004) Assessing the quality of evidence in evidence-based policy: why,how and when? ESRC Research Methods Programme Working Paper No 1.

Pawson, R. (2006a) *Evidence-Based Policy: A Realist Perspective*. London: Sage.

Pawson, R. (2006b) Realistic Evaluation: Seminar Tavistock Institute.

Pawson, R. (2013) *The Science of Evaluation: A Realist Manifesto* London: Sage.

Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*. London: Sage.

Pawson, R. and Tilley, N. (2004) *Realist Evaluation*. [online] Available from: http://www.communitymatters.com.au/RE_chapter.pdf (Accessed Dec).

Pawson, R. and Tilley, N. (2008) Realist evaluation. In: Otto, H.-U., Polutta, A. and Ziegler, H. eds. *Evidence-based Practice: Modernising the Knowledge Base of Social Work*. Leverkusen, Germany: Barbara Budrich: 151-184.

Phillips, D. C. (1987) *Philosophy, science, and social inquiry: Contemporary methodological controversies in social science and related applied fields of research*. Oxford, England: Pergamon Press.

Pintrich, P. R. (2000) An achievement goal theory perspective on issues in motivation terminology, theory and research. *Contemporary Educational Psychology*, 25: 92-104.

Pololi, L. H., Knight, S. M., Dennis, K. and Frankel, R. M. (2002) Helping medical school faculty realize their dreams: an innovative, collaborative mentoring program. *Academic Medicine*, 77 (5): 377-384.

Postareff, L., Lindblom-Ylanne, S. and Nevgi, A. (2007) The effect of pedagogical training on teaching in higher education. *Teaching and Teacher Education*, 23: 557-571.

Pratt, D. D. (1992) Conceptions of teaching. *Adult Education Quarterly*, 42 (4): 203 - 220.

Prebble, T., Hargraves, H., Leach, L., Naidoo, K., Suddaby, G. and Zepke, N. (2004) Impact of Student Support Services and Academic Development Programmes on Student Outcomes in Undergraduate Tertiary Study: A synthesis of the Research. *Ministry of Education*, New Zealand

Pretty, J., Guijt, I., Thompson, J. and Scones, I. (1995) *Participatory Learning and Action: A Trainer's Guide*. London: IIED.

Purcell, N. and Lloyd-Jones, G. (2003) Standards for medical educators. *Medical Education*, 37 (2): 149-154.

Putnam, H. (1999) *The threefold cord: Mind, body, and world*. New York: Columbia University Press.

Quality Assurance Agency for Higher Education (2000) *Handbook for Academic Review*. [online] Available from: http://www.qaa.ac.uk/Publications/InformationAndGuidance/ (Accessed Nov 2010).

Quine, S. (1985) Does the mode matter?' A comparison of three modes of questionnaire completion. *Community Health Studies*, 9: 151-156.

Rallis, S. F. and Rossman, G. B. (2003) Mixed methods in evaluation contexts: A pragmatic framweork. In: Tashakkori, A. and Teddlie, C. eds. *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: Sage: 491-512.

Ramani, S. (2006) Twelve tips to promote excellence in medical teaching. *Medical Teacher*, 28 (1): 19-23.

Reason, P. (1999) Integrating Action and Reflection through co-operative inquiry. *Management Learning*, 30 (2): 207-226.

Redfern, S., Christian, S. and Norman, I. (2003) Evaluating change in health care practice: lessons from three studies. *Journal of Evaluation in Clinical Practice*, 9 (2): 239-249.

Reeve, J. (2001) *Understanding motivation and emotion* 3rd. Fort Worth: Harcourt.

Reeve, J., Deci, E. L. and Ryan, R. M. (2004) Self-determination theory: A dialectical framework for understanding sociocultural influences on student motivation. In: McInerney, D. and Van Etten, S. eds. *Big theories revisited* Greenwich, CT: Information Age: 31-60.

Reid, A., Stritter, F. T. and Arndt, J. E. (1997) Assessment of faculty development program outcomes. *Family Medicine*, 29 (4): 242-247.

Rice, E. (2002) Beyond Scholarship Reconsidered: Toward an Enlarged Vision of the Scholarly Work of Faculty Members. *New Directions for Teaching and Learning*, 90: 7-17.

Richardson, J. T. E. (2005) Students' approaches to learning and teachers' approaches to teaching in Higher Education. *Educational Psychology*, 25 (6): 673-668.

Rider, E., Cooke, M. and Lowenstein, D. (2002) The Academies Collaborative: Sharing a New Model for Medical Education. *Academic Medicine*, 77 (5): 455.

Robb, D. and Heil, J. (2003) Mental causation. *Stanford Encyclopedia of Philosophy*.

Robins, L., Ambrozy, D. and Pinsky, L. (2006) Promoting Academic Excellence Through Leadership Development at the University of Washington: The Teaching Scholars Program. *Academic Medicine*, 81 (11): 979-983.

Robinson, T. and King, T. (2002) *Student perception of instructor affinity seeking, credibility and ethnicity in the HBCU classroom.* [online] Available from: http://www.xula.edu/xulanexus/issue4/Robinson_King.htm (Accessed 20th Sept 2013).

Robson, C. (2011) *Real World Research*. 3rd Edition. Chichester: Wiley.

Rosenbaum, M., Lenoch, S. and Ferguson, K. (2006) Increasing Departmental and College-Wide Faculty Development Opportunities Through a Teaching Scholars Program. *Academic Medicine*, 81 (11): 965-968.

Safman, R. M. and Sobal, J. (2004) Qualitative Sample Extensiveness in Health Education Research. *Health Education and Behaviour*, 31 (1): 9-21.

Saldaña, J. (1995) “Is Theatre Necessary?”: Final Exit Interviews with Sixth Grade Participants from the ASU Longitudinal Study. *Youth Theatre Journal*, 9 (1): 14-30.

Saldaña, J. (2013) *The coding manual for qualitative researchers*. London: Sage.

Saroyan, A. and Frenay, M. ed. (2010) *Building Teaching Capacities in Higher Education: A Comprehensive International Model*. Sterling, Virginia: Stylus.

Sayer, A. (1992) *Method in Social Science: A Realist Approach*. London: Routledge.

Sayer, A. (2000) *Realism and Social Science*. London: Sage.

Schaufeli, W. B., Martínez, I. M., Pinto, A. M., Salanova, M. and Bakker, A. B. (2002) Burnout and engagement in university students: A cross-national study. *Journal of Cross-Cultural Psychology*, 33 (5): 464-481.

Schimmel, K., Motley, D., Racic, S., Marco, G. and Eschenfelder, M. (2010) The importance of university web pages in selecting a higher education institution. *Research in Higher Education Journal*, 9 (10): 1-16.

Schmuck, R. A. (2006) *Practical Action Research for Change*. 2nd edition. Thousand Oaks, CA: Corwin Press.

Schofield, I. and Tolson, D. (2010) *Scottish Football Museum reminiscence pilot project for people with dementia: A realistic evaluation*. Glasgow: School of Health Glasgow Caledonian University.

Schön, D. (1987) *Educating the reflective practitioner*. San Francisco: Jossey-Bass.

Schön, D. A. (1983) *The reflective practitioner: how professionals think in action.* London: Temple Smith.

Schwarz, M. R. and Wojtczak, A. (2002) Global minimum essential requirements: a road towards competence-oriented medical education. *Medical Teacher*, 24 (2): 125-129.

Scott, D. (2000) *Realism and Educational Research: New perspectives and Possibilities*. London: Routlege.

Seabrook, M. (2001) Learning to teach. *Postgraduate Medical Journal*, 77 (908): 361-362.

Searle, N., Thompson, B. and Perkowski, L. (2006) Making It Work: The Evolution of a Medical Educational Fellowship Program. *Academic Medicine*, 81 (11): 984-989.

Sheets, K. J. (1985) Evaluating short-term training programs: a practical approach. *Journal of Instructional Development*, 8 (1): 26-32.

Sheets, K. J. and Henry, R. C. (1984) Assessing the impact of faculty development programs in medical education. *Journal of Medical Education*, 59 (9): 746-748.

Sheets, K. J. and Henry, R. C. (1988) Evaluation of a faculty development program for family physicians. *Medical Teacher*, 10 (1): 75-83.

Sheets, K. J. and Schwenk, T. L. (1990) Faculty development for family medicine educators: an agenda for future activities. *Teaching and Learning Medicine*, 2: 141-148.

Shepperd, S., Lewin, S., Straus, S., Clarke, M., Eccles, M. P., Fitzpatrick, R., Wong, G. and Sheikh, A. (2009) Can We Systematically Review Studies That Evaluate Complex Interventions? *PLOS Medicine*, 6 (8): e1000086.

Shulman, L. (1987) Knowledge and teaching: foundations of the new reform. *Harvard Education Review*, 57 (1): 1-22.

Silverman, D. (2006) *Interpreting Qualitative Data*. Third edition. London, UK: Sage Publications.

Silverman, D. (2010) *Doing Qualitative Research*. 3rd edition. London, UK: Sage.

Simpson, D., Marcdante, K., Morzinski, J., Meurer, L., McLaughlin, C., Lamb, G., Janik, T. and Currey, L. (2006) Fifteen Years of Aligning Faculty Development With Primary Care Clinician–Educator Roles and Academic Advancement at the Medical College of Wisconsin. *Academic Medicine*, 81 (11): 945-953.

Skeff, K. M. (1983) Evaluation of a method for improving the teaching performance of attending physicians. *American Journal of Medicine*, 75 (3): 465-470.

Skeff, K. M., Stratos, G., Mygdal, W., DeWitt, T. A., Manfred, L., Quirk, M., Roberts, K., Greenberg, L. and Bland, C. J. (1997a) Faculty development a resource for clinical teachers. *Journal of General Internal Medicine*, 12 (Supplement 2): S56 - S63.

Skeff, K. M., Stratos, G. A., Mygdal, W. K., DeWitt, T. A., Manfred, L., Quirk, M. E., Roberts, K. B. and Greenberg, L. W. (1997b) Clinical teaching improvement: past and future for faculty development. *Family Medicine*, 29 (4): 252-257.

Sobral, D. T. (2004) What kind of motivation drives medical students learning quests? *Medical Education*, 38 (9): 950-957.

Sorcinelli, M. D., Austin, A. E., Eddy, P. L. and Beach, A. L. (2006) *Creating the Future of Faculty Development: Learning from the Past, Understanding the Present*. San Francisco, CA: Jossey-Bass.

Sorinola, O. O. and Thistlethwaite, J. (2013) A systematic review of faculty development activities in family medicine. *Medical Teacher*, 35 (7): e1309-e1318.

Southern, N. L. (2007) Mentoring for transformative learning: the importance of relationship in creating learning communities of care. *Journal of Transformative Education*, 5 (4): 329-338.

Spady, W. G. (1988) Organising for results: the basis of authentic restructuring and reform. *Educational Leadership*, 46 (2): 4-8.

Stacey, R. D. (2010) *Strategic management and organisational dynamics: the challenge of complexity*. 6th ed. London: Financial Times / Prentice Hall.

Steinert, Y. (2000) Faculty development in the new millennium: key challenges and future directions. *Medical Teacher*, 22 (1): 44-50.

Steinert, Y. (2010a) Faculty development: From workshops to communities of practice. *Medical Teacher*, 32 (5): 425-428.

Steinert, Y. (2010b) Making It All Happen: Faculty Development for Busy Teachers. In: Peter, C. and Diana, W. eds. *ABC of Learning and Teaching in Medicine*. London: Wiley-Blackwell: 73-77.

Steinert, Y. (2012) Perspectives on faculty development: aiming for 6/6 by 2020. *Perspectives on Medical Education*, 1 (1): 31-42.

Steinert, Y. (2013) Staff development. In: Dent, J. A. and Harden, R. M. eds. *A Practical Guide for Medical Teacher* London: Elsevier, Churchill, Livingston: 369-376.

Steinert, Y., Nasmith, L., Daigle, N. and Franco, E. (2001) Improving teacher's skills in working with 'problem' residents: a workshop description and evaluation. *Medical Teacher*, 23 (3): 284-288.

Steinert, Y., Nasmith, L., McLeod, P. and Conochie, L. (2003) A teaching scholars program to develop leaders in medical education. *Academic Medicine*, 78 (2): 142-149.

Steinert, Y., Cruess, S., Cruess, R. and Snell, L. (2005) Faculty development for teaching and evaluating professionalism: from programme design to curriculum change. *Medical Education*, 39 (2): 127-136.

Steinert, Y., Mann, K., Centeno, A., Dolmans, D., Spencer, J., Gelula, M. and Prideaux, D. (2006) A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education: BEME Guide No. 8. *Medical Teacher*, 28 (6): 497-526.

Steinert, Y. and McLeod, P. (2006) From Novice to Informed Educator: The Teaching Scholars Program for Educators in the Health Sciences. *Academic Medicine*, 81 (11): 969-974.

Steinert, Y., McLeod, P., Boillat, M., Meterissian, S., Elizov, M. and Macdonald, M. E. (2009) Faculty devevelopment: a 'Field of dreams'? *Medical Education*, 43 (1): 42-49.

Steinert, Y., Macdonald, M. E., Boillat, M., Elizov, M., Meterissian, S., Razack, S., Ouellet, M.-N. and McLeod, P. J. (2010) Faculty development: if you build it, they will come. *Medical Education*, 44 (9): 900-907.

Steinert, Y., Naismith, L. and Mann, K. (2012) Faculty development initiatives designed to promote leadership in medical education. A BEME systematic review: BEME Guide No. 19. *Medical Teacher*, 34 (6): 483-503.

Stephenson, A., Todres, M. and Jones, R. (2009) Reply to Dornan et al.'s 'On evidence'. *Medical Education*, 43 (4): 390-391.

Stephenson, J. ed. (2001) *Teaching & Learning Online: New Pedagogies for New Technologies*. Abingdon, Oxon: Routledge Falmer.

Streiner, D. L. and Norman, G. R. (2008) *Health Measurement Scales; a Practical Guide to their Development and Use*. NY: Oxford.

Stritter, F. T. (1983) Faculty evaluation and development. In: McGuire, C. H., Foley, R. P., Gorr, A. and Richards, R. W. eds. *Handbook of Health Professions Education*. San Francisco, CA: Jossey-Bass: 294-318.

Stronge, J. H., Ward, T. J. and Grant, L. W. (2011) What Makes Good Teachers Good? A Cross-Case Analysis of the Connection Between Teacher Effectiveness and Student Achievement. *Journal of Teacher Education*, 62 (4): 339-355.

Swain, H. (2009) 'A hotchpotch of subjectivity'. *Guardian, London.* Date: 19.05.2009.

Swain, J. (2004) The resources and strategies that 10-11-year-old boys use to construct masculinities in the school setting. *British Educational Research Journal*, 30 (1): 167-185.

Swanson, L. A. (2009) Value Perceptions as Influences upon Engagement. *Innovative Higher Education*, 34 (4): 269-281.

Taylor, E. W. (1998) *The Theory and Practice of Transformative Learning: A Critical Review*. Information Series no 374. Columbus: ERIC Clearinghouse on Adult, Career, and Vocational Education, Center on Education and Training for Employment, College of Education, the Ohio State University.

Teherani, A., Hitchcock, M. and Nyquist, J. (2001) Longitudinal Outcomes of an Executive-model Program for Faculty Development. *Academic Medicine*, 76 (10 Suppl): S68-S70.

Tekian, A. and Harris, I. (2012) Preparing health professions education leaders worldwide: A description of masters-level programs. *Medical Teacher*, 34 (1): 52-58.

Thangaratinam, S., Barnfield, G., Weinbrenner, S., Meyerrose, B., Arvanitis, T., Horvath, A., Zanrei, G., Kunz, R., Suter, K., Walczak, J., Kaleta, A., Rengerink, K., Gee, H., Mol, B. and Khan, K. (2009) Teaching trainers to incorporate evidence-based medicine (EBM) teaching in clinical practice: the EU-EBM project. *BMC Medical Education*, 9 (1): 59.

Thibault, G., Neill, J. and Lowenstein, D. (2003) The Academy at Harvard Medical School: Nurturing Teaching and Stimulating Innovation. *Academic Medicine*, 78 (7): 673-681.

Tilley, N. (1993) *Understanding Car Parks, Crime and CCTV, Crime Prevention Unit Paper 42*. London: The Home Office.

Timmins, P. and Miller, C. (2007) Making evaluations realistic: the challenge of complexity. *Support for Learning*, 22 (1): 9-16.

Todres, M., Stephenson, A. and Jones, R. (2007) Medical education research remains the poor relation. *British Medical Journal*, 335 (7615): 333-335.

Tolson, D., McIntosh, J., Loftus, L. and Cormie, P. (2007) Developing a managed clinical network in palliative care: a realistic evaluation. *International Journal of Nursing Studies*, 44 (2): 183-195.

Torr, A. (2008) A complex view of professional competence.

Towle, A. (1991) *Critical Thinking. The Future of Undergraduate Medical Education*.

Trinder, J. C. (2008) Competency Standards - A measure of the quality of a workforce. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Conference*.

Vereniging van Universiteiten (2008) *Basiskwalificatie Onderwijs. (Association of Universities in the Netherlands. Basic Qualification in Teaching) .* . [online] Available from: http://www.vsnu.nl/Beleidsterreinen/Lopende-dossiers/Basiskwalificatie-Onderwijs.htm (Accessed 30 Mar 2012).

Vroom, V. H. (1964) *Work and Motivation*. New York: Wiley.

Waldman, J. D., Kelly, F., Arora, S. and Smith, H. L. (2004) The shocking cost of turnover in health care. *Health Care Management Review*, 29 (1): 2-7.

Weberschock, T., Sorinola, O., Thangaratinam, S., Oude Rengerink, K., Arvanitis, T. N., Khan, K. S. and Group*, E. U. (2013) How to confidently teach EBM on foot: development and evaluation of a web-based e-learning course. *Evidence Based Medicine*, 18 (5): 170-172.

Weiner, B. (1974) *Achievement motivation and attribution theory*. Morristown, NJ: General Learning Corporation.

Weiss, R. S. (1994) *Learning from strangers: The art and method of qualitative interviewing*. New York: The Free Press.

Weiten, W. (2013) *Psychology: Themes and Variations*. 9th edition. Belmont, CA: Wadsworth Cengage Learning.

WFME (2000) WFME Task force on defining international standards in basic medical education. Report of the working party. *Medical Education*, 34 665-675.

Whitcomb, M. (2002) Competency-based Graduate Medical Education? Of Course! But How Should Competency Be Assessed? *Academic Medicine*, 77 (5): 359-360.

Whitcomb, M. (2003) The Medical School’s Faculty Is Its Most Important Asset. *Academic Medicine*, 78 (2): 117-118.

Wilkerson, L. and Irby, D. (1998) Strategies for improving teaching practices: a comprehensive approach to faculty development. *Academic Medicine*, 73 (4): 387-396.

Wilkerson, L., Uijtdehaage, S. and Relan, A. (2006) Increasing the Pool of Educational Leaders for UCLA. *Academic Medicine*, 81 (11): 954-958.

Wilkinson, T. J., Wells, J. E. and Bushnell, J. A. (2007) Medical student characteristics associated with time in study: Is spending more time always a good thing? *Medical Teacher*, 29 (2-3): 106 -110.

Wingard, D. L., Garman, K. A. and Reznik, V. (2004) Facilitating faculty success: outcomes and cost benefit of the UCSD National Center of Leadership in Academic Medicine. *Academic Medicine*, 79 (10): S9-11.

Wojtczak, A. and Schwarz, M. R. (2000) Minimum essential requirements and standards of medical education. *Medical Teacher*, 22 (6): 555-559.

Wolcott , H. F. (1990) On seeking and rejecting validity in qualitative research. In: Eisner, E. W. and Peshkin, A. eds. *Qualitative Inquiry in Education: The Continuing Debate*. New York: Teachers College Press: 121-152.

Wong, G., Greenhalgh, T., Westhorp, G. and Pawson, R. (2012) Realist methods in medical education research: what are they and what can they contribute? *Medical Education*, 46 (1): 89-96.

World Health Organization (2013) Transforming and scaling up health professionals' education and training.

Wright, S. (1921) Correlation and causation. *Journal of Agricultural Research*, 20 (7): 557-585.

Yardley, S. and Dornan, T. (2012) Kirkpatrick's levels and education 'evidence'. *Medical Education*, 46 (1): 97-106.

Yonezawa, S., Jones, M. and Joselowsky, F. (2009) Youth engagement in high schools: Developing a multidimensional critical approach to improving engagement for all students. *Journal of Educational Change*, 10 (2): 191-209.

Zibrowski, E., Weston, W. W. and Goldszmidt, M. (2008) 'I don't have time': issues of fragmentation, prioritisation and motivation for education scholarship among medical faculty. *Medical Education*, 42 (9): 872-878.

Zimmerman, B., Lindberg, C. and Plsek, P. (1998) *Edgeware: Lessons from Complexity Science for Health Care Leaders*. Irving Texas: VHA, Inc.

# Appendix 1

# List of Medical Education Journals

Academic Medicine

Advances in Health Sciences Education

BMC Medical Education

British Medical Journal

Clinical Teacher

Journal of Continuing Education in the Health Professions

Journal of the American Medical Association

Lancet

Medical Education

Medical Teacher

Teaching and Learning in Medicine

# Appendix 2
# FD as defined by various authors over the decades

| Learning Theories | Authors | Definitions / Descriptions |
|---|---|---|
| *Behavioural Theory* | Gaff (1975) | Faculty development are activities that assist teachers to improve their teaching skills, design better curricula, and / or improve the institutional climate for education |
| | Francis (1975) | Defined FD as an institutional process which seeks to modify the attitudes, skills, and behaviour of faculty members toward greater competence and effectiveness in meeting student needs, their own needs, and the needs of the institution. |
| | Centra (1976) | FD is a broad range of activities institutions use to *renew* or *assist* faculty in their varied roles |
| *Cognitive Theory* | Nelson (1983) | FD refers to any endeavour designed to improve faculty performance in all aspects of their professional lives – as scholars, advisers, academic leaders, and contributors to institutional decisions |
| | Stritter (1983) | No precise definition but divided FD into three categories: technical assistance (more or less at an individual level), high faculty involvement (e.g. through workshops; collaborative educational research) and assessment (by peers, students and self-assessment, with feedback). |
| | Bland & Schmitz (1988) | Defined FD as providing faculty and institutional vitality, and suggesting strategies to improve vitality at 3 levels: institution (e.g. altering personnel policies, redefining mission), department (organisational development and practice, e.g. providing administrative assistance), individual faculty members (faculty exchange, peer consultation, cross-departmental teaching) |
| *Social Learning Theory* | Bland et al. (1990) | FD is a planned programme to *prepare* institutions and faculty members for their academic roles and to improve an individual's knowledge and skills in the areas of teaching, research and administration |
| | Sheets & Schwenk (1990) | FD is any planned activity to improve an individual's knowledge and skills in areas considered essential to the performance of a faculty member in a department or a residency programme (e.g. teaching, administrative, research and clinical skills). |
| | Hitchcock et al. (1993) | Major conclusions from a review of the literature: Concept of FD is evolving and expanding (e.g. professional academic skills; ethics, clinical and research skills) and Teaching skills are still a prominent aspect of FD. Institutional environment is important in FD (to improve productivity). Faculty evaluation is an effective approach to FD but more research into outcomes of FD is required |
| | Lewis (1996) | Used the term FD to denote all the activities and programmes that are typically used in the improvement of teaching and learning in higher education |
| | Wilkerson & Irby (1998) | Defined FD as a tool for improving the educational vitality of institutions through attention to the competencies needed by individual teachers and to the institutional policies required to promote academic excellence |
| *Outcome Based Education (OBE)* | Steinert (2000) | Explained that to keep pace with changes, faculty development will need to broaden its focus by using diverse learning methods, underpinned by learning theories, fostering partnerships / collaboration, and rigorously evaluating interventions |
| | Whitcomb (2003) | Suggests that the goal of FD is to teach faculty members the skills relevant to their institutional setting & faculty position,& to sustain their vitality both now & in the future |
| | Bligh (2005) | Faculty development programmes are outward signs of the inner faith that institutions have in their workforce |
| | Harris et al. (2007) | Tried to categorise faculty members as follows: Teacher-Administrator (e.g. Chair; Residency Director; Clinic Director); Teacher-Educator (e.g. Director of Education; Pre-doctoral Director; Clerkship Director); Teacher-Researcher / Teacher-Clinician |
| | McLean et al (2008) | Defined FD at institutional level as the personal and professional development of teachers, clinicians, researchers and administrators to meet the goals, vision and mission of the institution in terms of its social & moral responsibility to the communities it serves |

# Appendix 3

## Appendix 3a. Ethics approval

13 October 2011

Warwick Medical School

PRIVATE
Mr Olanrewaju Sorinola
Obstetrics and Gynaecology Department
South Warwickshire Foundation NHS Trust
Warwick Hospital
Lakin Road, Warwick
CV34 5BW

Dear Lanre,

Study Title and BREC Reference: *Faculty Development Activities in UK Medical Schools – 149-09-2011*

Thank you for submitting your revisions to the above-named project to the University of Warwick Biomedical Research Ethics Sub-Committee for Chair's Approval.

I am pleased to confirm that I am satisfied that you have met all of the conditions and your application meets the required standard, which means that full approval is granted and your study may commence.

I take this opportunity to wish you success with the study and to remind you any substantial amendments require approval from the committee before they can be made. Please keep a copy of the signed version of this letter with your study documentation. The committee would also welcome an End of Project Report.

Yours sincerely,

Jane Barlow

Jane Barlow
Chair
Biomedical Research
Ethics Sub-Committee

Copy:
File
Dr. David Davies

**Biomedical Research Ethics Subcommittee**
Enquiries: Clair Henrywood
Tel: 02476-528207
Email: brec@warwick.ac.uk

THE UNIVERSITY OF WARWICK

# Appendix 3b. Participant information for observation and interview

**Participant Information for Observation and Interview of Participants Attending Faculty Development Activity**

## Study Title

A realist evaluation of faculty development in UK medical schools

## Summary

This research is an in depth exploration of medical school educators' perceptions of faculty development (FD) activities designed to improve teaching effectiveness in UK medical schools. Doctors are experts in what they teach but most have had little or no training in how to teach, hence the need for FD. Currently, the extent, scope and views of FD activities in the UK medical schools remain uncharted and it's uncertain whether FD achieves its stated objectives. The study is in three phases. In phase I, a scoring index was used to score the faculty development webpages of the 33 UK medical schools. In phase II, I plan to interview the FD coordinator at Warwick Medical School and observe two faculty development courses. I would like to interview participants (non-taped) during the course and carry out one-hour, taped interviews of six consenting participants from each course six months later. This will be followed in phase III by one hour taped interviews of one medical educator and one FD coordinator from eight UK medical schools. This study will enable me to ascertain views of medical educators' on the determinants of effective faculty development activities. I will report perceptions on whether FD achieves the purpose of improving teaching effectiveness, and establish what works (or doesn't) in FD, in what context, for whom and why. The results can inform future faculty development and I will recommend strategies to enhance acceptability, efficacy and utility of FD.

## Invitation

I would like to invite you to take part in our research study. Before you decide I would like you to understand why the research is being done and what your involvement will be. Please take time to read the following information carefully. I would estimate that this should take about 10 minutes. Feel free to talk to others about the study if you wish. Do ask if there is anything that is not clear or if you would like more information. My details are at the end of this sheet.

## What is the purpose of the study?

In the UK, medical students are taught by clinicians with varying educational knowledge hence the need for faculty development to improve educational skills. However despite the time and resources devoted to faculty development, the views of medical educators on its usefulness have never been ascertained. The aim of this study is to get the views of both faculty development providers and the participants on acceptability, efficacy and utility of FD activities aimed at teaching improvement. This will help to tailor faculty development to medical educators needs and maximise the gains from such activities.

## Why have I been invited?

You are being invited to take part in the study because you are participating in the three days 'Essentials in Clinical Education' course at Warwick Medical School (WMS). This research will involve group observation as well as individual observation and interview of participants during the course. Depending on the responses obtained, four participants from each course will be invited for a follow-up interview six months later.

## Do I have to take part?

As the nature of the research study is entirely voluntary, it is up to you to decide. We will describe the study and go through this information sheet with you. If you agree to take part, we will then ask you to sign the consent forms. There are separate consent forms for observation and interview as some people might wish to be observed but not interviewed or vice versa. In addition, in case there are session(s) that you do not want observed or if after a particular session (e.g. involving an emotive or personal discussion) you want the data removed, this will be done. You are free to withdraw at any time, without giving a reason. Additionally, if you agree to take part and then decide to withdraw at a later date, you will be able to do this and any information personal to you can be removed from the study report if you so choose.

## What will happen to me if I take part?

As a participant, you will be observed as part of the group and as an individual during sessions, discussions, activities, breaks and interviewed at your convenience. Group observation data will include session organisation, number of attendees, topics covered, teaching methods, evaluation method(s), discussions during sessions and breaks while individual's observation will include engagement, interaction, discussion, question and answers. The interviews during the course will not be taped, will last about 10 minutes each

and will happen three times in total over the three day period. Interviews will seek your opinions on the usefulness, relevance and applicability of the faculty development session(s) you have participated in. Using the data obtained, six participants from each course will be invited at their convenience for a follow up interview six months later. This audio-taped interview will last for one hour and focus on the impact of the educational intervention i.e. what has changed. The interview will also ask for some personal information to see if there is any relationship between aspects such as knowledge, educational qualifications and gains from faculty development. However, all of the information will be kept confidential and will only be used for the study.

## What are the possible disadvantages and risks of taking part?

1. During the observation and interviews, participants are very unlikely to disclose any sensitive personal information or be asked sensitive personal questions. However, if the participant discloses any personal information or issue which (s)he does not feel comfortable to be included in the study, it will be excluded.
2. It is most unlikely that you will be inconvenienced by taking part in the observation or interview during the course. However, I do appreciate it may become inconvenient to participate in the six months interview due to clinical commitments. To minimise this, the interview will be conducted at your convenience.

## What are the possible benefits of taking part?

The benefits of this study are many. Interviewing medical educators is important as they experience the training and in turn are expected to deliver teaching; hence ascertaining their views is an important first step in assessing the scope for quality improvement. Participants will be able to express their views and reflect on the educational values of these interventions. It will be useful as evidence of reflective practice for your teaching portfolio which can be used at appraisals. In addition, some people will find the experience interesting and revealing for their own educational development. The information we get from this study will also benefit future faculty development organisers, and the profession to receive a return through improved teaching and training for the next generation of doctors.

## What if I want more information about the study?

If you have any questions about any aspect of the study or your participation in it please contact me. My details are at the end of this sheet.

## What if there is a problem?

If you have any concerns about any aspect of this project, you should speak to me and I will do my best to answer your questions. However, if you remain unhappy or feel uncomfortable in doing this you can contact David Davies, my academic supervisor at the University of Warwick. His contact details are: david.davies@warwick.ac.uk. Tel: 02476150192

## Who should I contact if I wish to make a complaint?

Any complaint about the way you have been dealt with during the study or any possible harm you might have suffered will be addressed. Please address your complaint to the person below who is a senior University official entirely independent of the study:

Nicola Owen
Deputy Registrar
Deputy Registrar's Office
University of Warwick
Coventry,
CV4 8UW
Tel: 024 7652 2713
E-mail: Nicola.Owen@warwick.ac.uk

## Will my taking part in the study be kept confidential?

Yes. All information which is collected about you during the course of the research will be kept strictly confidential, and any information about you will have your name and address removed so that you cannot be recognised. During the study data will be stored on a non-shared password protected file on a university computer and will be accessed only by the investigator. After the study the data will be kept for seven years after which it will be destroyed. It will not be possible to identify you from any published material arising from the study.

## What will happen if I don't want to carry on with the study?

During the course of the research study, you can withdraw at any time. If you withdraw, the option will be given of whether you prefer your data to be included in the study or excluded.

## What will happen to the results of the research study?

The research findings will be presented locally to educators, health care professionals, medical education department(s), as well as national and international conferences. The results may also be disseminated through peer reviewed journals; anonymity of participants will be maintained at all times.

## Who is organising and funding the research?

It is being undertaken as a part of a PhD study and sponsored by the University of Warwick.

## Who has reviewed the study?

All research is looked at by an independent group of people, called the Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the University of Warwick's Biomedical Research Ethics Committee.

## Further information and contact details

For further information regarding the study, you can contact the chief investigator:

Lanre Sorinola Contact number: 07753749488

Email address: olanrewaju.o.sorinola@warwick.ac.uk

Thank you for taking the time to read this.

## Appendix 3c. Higher Education Academy Descriptors

Descriptor 1: An Associate Fellow of the Academy (demonstrates an understanding of specific aspects of effective teaching, learning support methods and student learning).

Descriptor 2: A Fellow of the Academy (demonstrates a broad understanding of effective approaches to teaching and learning support as key contributions to high quality student learning).

Descriptor 3: A Senior Fellow of the Academy (demonstrates a thorough understanding of effective approaches to teaching and learning support as a key contribution to high quality student learning).

Descriptor 4: A Principal Fellow of the Academy (demonstrates a sustained record of effective strategic leadership in academic practice and academic development as a key contribution to high quality student learning).

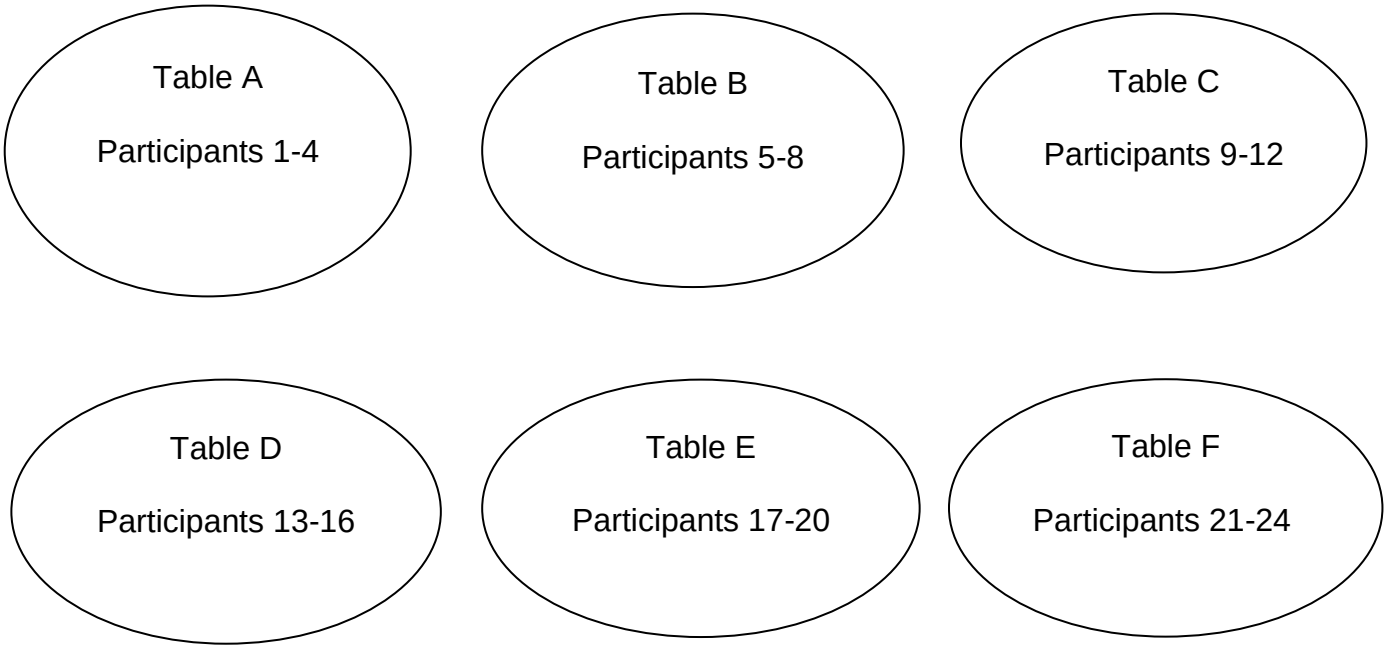
## Appendix 3d. Phase 2 observation schedule

**Session Details:** Date: Time:

Session Topic: Facilitator**:**

Learning Outcomes: Methods used in session:

**Seating Arrangement / Coding**




**Codes:** Jan cohort: 1J – 24J; Apr cohort: 1A – 24A

| **A. Group Observations** | **B. Individual Observations** |
|---|---|
| **C. Brief Summary** | |

# Appendix 3e. Phase II interview schedule

## Interview Schedule for Course Participants – 6 months

### Introduction

Thank you for your interest in participating in this interview for my research on FD. This should take about an hour. I will start by explaining my hypotheses and the general theoretical ground that I am exploring so that you have a clear idea of the concepts that I wish to discuss. We will then discuss these hypotheses using information from the FD course as well as information from your experience of teaching during the six months following the course.

(Choose questions from domains below as guide and explore as appropriate)

### Q1. Course Content / Process

1. Tell me your thoughts about the course now that 6 months has passed.
2. Have you completed the course or not? If not why not. Explore what a participant means by successful
3. What is your view on the organisation of the course? – Registration, online materials, assignments.

### Q2. Motivation / Attendance

4. What was your reason for attending the course? Has this been satisfied?
5. In which areas of educational skills or knowledge did you have the greatest gaps and were your needs met during the course?
6. Have you kept in touch with any of the other participants of the course?

### Q3. Engagement

7. Describe your engagement during the course
8. Tell me about the assessment. Explore the impact of the assessment – Was it necessary? Did it drive your learning?
9. What would you have liked included in the course? What was missing at the time and now on reflection what else would have been helpful?

### Q4. Perception – Usefulness / Relevance

10. How useful was the course you attended in improving your teaching practice?
11. How relevant have you found the areas covered in the course for your own practice?
12. Have you been able to apply any of the skills learnt? Give examples
13. In your view who is the client in faculty development? *Participants' students, faculty member or the organisation.*

### Q5. Impact / Outcomes

14. What has been the impact of the course on your teaching practice i.e. what has changed? And in other ways on yourself – so could be learning in areas not directly applicable to own teaching practice but to 'personal development'
15. What do you do differently now and why?
16. What have you accomplished (or opportunity presented to you) that would not have happened without this training?
17. In your opinion, what has been the greatest impact of the course so far? *Probe: For yourself, your learners, your organisation*
18. Any change in the organisation of teaching in a broader sense rather than personal? *Depending on the participant's role: (e.g. people who have been on similar courses have gone back and ensured that learning outcomes are defined for all programmes and become involved in the development of their colleagues)*
19. Any further plan for development / scholarship? Career plans – Research / publications / further qualifications?

### Q6. Views on FD / Future

20. What do you think of faculty development in general (i.e. being taught how to teach)?
21. What should be the focus of faculty development activities?
22. If you have a choice, what type(s) of faculty development activity will you choose and why? – Workshops (½ - 1 day), short courses (3-5 days), seminar series, longitudinal programmes (1-2 years). Are certificates / accreditation important?
23. Why do you think some teachers do not participate in faculty development?
24. Your views on the future of FD / Any other comment

# Appendix 4

## 4a. Interview schedule for medical educators

### Introduction

Thank you for your interest in participating in this interview for my research on FD. This should take about an hour. I will start by explaining my hypotheses and the general theoretical ground that I am exploring so that you have a clear idea of the concepts that I wish to discuss. We will then discuss these hypotheses using information from the FD course as well as information from your experience of teaching during the six months following the course.

### Demographic

Name, Title / Role: FT / PT

Years at this institution

Educational Background (brief)

(Choose questions from domains below as guide and explore as appropriate)

### Explore FD realist hypotheses

1. Can you tell me about faculty development activities in your school?
2. What types of FD activities are available? *Probe – what are the key areas covered in FD*
3. When did you last attend / participate in a faculty development activity to improve your educational skills? *Probe – voluntary, mandatory, promotion linked.*
4. Which types of FD activities do you engage in and why? *Probe – what motivates, what engages, etc.*
5. Which aspect of FD has produced the greatest impact and why? *Why do you think the programme works or not.*
6. How useful was the FD you attended in improving your teaching practice? Or how relevant to your own practice? *Probe – value of FD*
7. In your view who is the client in FD? *Probe – participants' students, faculty member or the organisation.*
8. How important do you believe teaching is to your own professional practice? *Probe – how often do you discuss teaching with colleagues, practice self-assessment,*

*invite peer feedback /observation, time spent preparing lectures, tutorials.*

9. How important do you believe teaching is to your institution in judging your accomplishments? *(Probe – annual reviews, promotion / tenure, merit awards)*

## Outcomes / Impact

10. What has been the impact of the FD activity on your teaching practice? *What has changed?*
11. In your opinion, what has been the most important impact of FD? *Probe – personal, professional.*
12. What have you accomplished (or opportunity presented to you) that would not have happened without this training? *i.e. might not have been able to do without FD training*
13. Describe career development / changes in the last 2 years?
14. In your view how effective are the different types of FD activities provided by your medical school.

## Views on FD / Future

15. To what extent do you feel your faculty or medical school support your growth as a tutor?
16. Are there any barriers to FD in your institution? *If so what are they?* Why do you think some teachers do not participate in faculty development?
17. If you have a choice of participating in several FD programmes with similar content, what aspects of delivery would cause you to register for one over the other? *Probe – workshops (½ - 1 day), short courses (3-5 days), seminar series, longitudinal programmes (1-2 years), certificates / accreditation?*
18. Your views on the future of FD / Any other comment

## 4b. Interview schedule for faculty development coordinators

### Introduction

Thank you for your interest in participating in this interview for my research on FD. This should take about an hour. I will start by explaining my hypotheses and the general theoretical ground that I am exploring so that you have a clear idea of the concepts that I wish to discuss. We will then discuss these hypotheses using information from the FD course as well as information from your experience of teaching during the six months following the course.

### Demographic

Name, Title / Role: FT / PT.
How long have you held a position of responsibility in FD?
Years (Total) / Years at this institution
Educational Background (brief)

### Explore FD set up – Generic / Governance

1. Can you tell me about faculty development at this school? *Overall goal, benefits*
2. How is FD organised at this school? – *Is there a responsible committee or individual?*
3. What types of FD activities are available? How is each quality assured? *Show FD chart*
4. What are the key areas covered in FD in your school? *Topic areas: Education, Research, Management, Personal. Ask for the list for the academic year.*
5. Do you keep records? If so, how many have attended FD activities over the last 3 years? *Increasing or decreasing - reasons*
6. Explore influences (internal / external). Is any faculty development activity compulsory – planning to make it? *Stipulated by whom?* What are the key challenges / barriers to faculty development in your institution? Explore professionalization
7. Explore resources – What is the source of funding, support, personnel? *Direct vs. Indirect funding.*

### Explore FD realist hypotheses

8. Which theoretical framework do you use in designing your programmes? – *content, delivery, assignments.*

9. Explore motivation – How do you encourage participation in faculty development? *Needs assessment, ownership, participation, feedback*
10. Which areas of faculty skills, knowledge and understanding have the greatest gaps and how were they identified? Have you conducted a systematic needs assessment of your faculty? *How long ago*
11. Explore engagement – What methods are used in the deliveries of the courses and programmes? Has anything changed because of feedback? *Learners, Peers*
12. In your view who is the client in faculty development? *Probe – students, faculty member or the organisation.*
13. If you could change the way you currently design, deliver or manage FD activities, what would they be?
14. Why do you think the FD programme works or not?

**Outcomes / Impact**

15. Does FD on teaching achieve its aim? *On what basis?*
16. In your opinion what has been the greatest impact of FD so far? *(Probe – for your learners, organisation)*
17. What aspect of the program has produced the greatest impact? *Probe why do you think the program works*
18. What methods of evaluation / outcomes are used in your programmes, courses, activities? *Why those methods.* How effective are the different types of FD activities you provide? *Give reasons*
19. How important do you believe teaching is to your institution in judging your educators accomplishments? (Probe – annual reviews, promotion, merit awards)

**Views on FD / Future**

20. In what direction do you think the field of faculty development should move in the next five years
21. Any other comment on FD or the future of FD

# Appendix 5

# Summary of Unengaged Phase II participants

**Participant 7J**

7J is a clinical education lead at a local hospice and runs staff development study days for nurses and doctors. He/she already has a PGCert in education and was attending the ECE as an optional module for Masters in Palliative Care that he/she was doing. The course was funded by the participants' organisation. Observation over the three days showed a fleeting engagement. There were times that he/she contributed in his/her group but only on one occasion over the three day period did he/she took the lead in explaining the strategy on behalf of the group during a group exercise. Majority of the time he/she appeared bored and seemed to need more depth on the topic to activate his/ her interest. I noted this in my observer record. The participant later re-affirmed this during the interviews and described the sessions as repetitive and confirmed that he/she wanted more depth in the areas covered.

Interestingly at the 6 months follow up interview, participant 7J who was consistently the most unengaged, felt on reflection he/she was a bit too judgemental because having already done the PGCE, he/she found most of the sessions repetitive but still learnt some useful tips. He/she has been able to utilise some of the learning and incorporate changes in delivering education. He/she gave examples of using a more interactive teaching approach and a more collaborative approach in designing programmes by checking what learners needs were and involving them in teaching. So why it might seem he/she was unengaged, it suggests some learning took place.

**Participant 24A**

Participant 24A is a specialist trainee near the end of the training programme. He/she teaches 4th year medical student during their six weeks block (does weekly case based discussion on various topics). The interesting thing is that the course was self-funded and one of his/her key motivation for attending was to know whether his/her teaching was effective and if he/she is teaching correctly. He/she was keen to develop a variety of skills and techniques to be able to teach and adapt to different levels of teaching. The other motivation was to advance his/her career as he/she was not interested in research and teaching was going to be his/her 'tick box'. According to the participant, "teaching has opened up a new avenue other than research to get your brownie points because not everybody is research minded and people want different avenues to pursue and teaching is a great one".

Observation over the three day period showed fluctuating engagement. He/she seemed to enjoy the discussions and interactivity especially during the sessions on 'on the job teaching' and 'teaching and learning in large groups' but overall had difficulty concentrating for long periods. There were times when the participant placed head on the table and was drowsy especially during the afternoon sessions. It might be possible that the previous attendance at a Teaching Improvement Programme (TIP) course might be a mitigating factor as some of the topics would have previously been covered which might explain the peaks and dips in engagement. The participant re-affirmed during the course interviews that he / she found some sessions too long and uninteresting hence he/she switches off.

# Appendix 6
# Publications

1. **Sorinola O**. Thistlethwaite J, David Davies, Ed Peile. Faculty development for educators: a realist evaluation. *Advances in Health Sciences Education*. 2014; 19 (3): 1-19
2. **Sorinola O**. Thistlethwaite J, David Davies. Motivation to engage in personal development of the educator. *Education for Primary Care*. 2013; 24(4): 226 – 229
3. **Sorinola O**, Gerzina T, Thistlethwaite J. Health Professional Education Programs; How the Teacher Develops. In: S. Loftus, M. Smith, J. Higgs, T. Gerzina, E. Duffy, (Eds). *Educating Health Professionals: Becoming a university teacher*. Sense Publishers. 2013: Vol 8; 49 – 60
4. **Sorinola O**, Thistlethwaite J. A systematic review of faculty development activities in family medicine. *Medical Teacher*. 2013, e1–e10, Early Online.